VOLUME IV MARCH, 1906 No. 4

TABLES FOR FACILITATING THE COMPUTATION OF PROBABLE ERRORS.

BY WINIFRED GIBSON, B.Sc., Jessel Student, University College, London.

THE following tables have been prepared with the view of rendering the computation of the probable errors of the mean, the standard deviation and the coefficient of variation as brief as possible. It is well known that if σ be the standard deviation, V the coefficient of variation, and n the number of observations, then:

(i) The Probable Error of the Mean $= \cdot 6744898\, \sigma/\sqrt{n}$.

(ii) The Probable Error of the Standard Deviation $= \cdot 6744898\, \sigma/\sqrt{(2n)}$.

(iii) The Probable Error of the Coefficient of Variation

$$= \cdot 6744898\, V\left\{1 + 2\left(\frac{V}{100}\right)^2\right\}^{\frac{1}{2}} \Big/ \sqrt{(2n)}.$$

It is clear, therefore, that the tabling of the expressions $\cdot 6744898/\sqrt{n}$, $\cdot 6744898/\sqrt{(2n)}$, and $V\left\{1 + 2\left(\frac{V}{100}\right)^2\right\}^{\frac{1}{2}}$ would make the computations of these probable errors much easier.

The question then arose whether the actual values or the logarithms of these expressions should be given. To those who work with a calculator, the latter would be of little service, while the former would still be helpful to those who have not this advantage. As nearly all biometricians now use a Tate, Brunsviga, or other mechanical computator, it was settled that the actual values should be given.

With the numbers usually available in biometric work, exclusive of human vital statistics, n rarely exceeds 1000. Hence

$$\chi_1 = \cdot 6744898/\sqrt{n} \quad \text{and} \quad \chi_2 = \cdot 6744898/\sqrt{(2n)}$$

have been tabled for $n = 1$ to 1000. Clearly for any even value of n between 1000 and 2000 the value of χ_1 can be found from the value of χ_2 for $\frac{1}{2}n$. Thus χ_1 for $n = 1594$ is χ_2 for $n = 797$, and is found to be ·01689. As the first difference for $\chi_2 > 500$ does not exceed ·00002, and the second difference is zero to five figures, it will be clear that χ_1 can for odd numbers between 1000 and 2000 be found by the simple process of halving the sum of the adjacent values of χ_2 for $\frac{1}{2}(n-1)$ and $\frac{1}{2}(n+1)$. Further, for any even value of n between 1000 and 2000, χ_2 may be found by halving the value of χ_1 for $\frac{1}{2}n$, and for any odd value of n between 1000 and 2000 χ_2 may be found by quartering the sum of the values of χ_1 for $\frac{1}{2}(n-1)$ and $\frac{1}{2}(n+1)$. Thus suppose $n = 1731$; we have $\frac{1}{2}(n-1) = 865$ and $\frac{1}{2}(n+1) = 866$. Hence $\chi_1 = \frac{1}{2}(\cdot01622 + \cdot01621) = \cdot016215$, and $\chi_2 = \frac{1}{4}(\cdot02293 + \cdot02292) = \cdot0114625$. The actual values are $\chi_1 = \cdot016212$ and $\chi_2 = \cdot011463$. Such differences will hardly ever have any statistical importance. To determine the probable error of the coefficient of variation we use Table III. Here we find

$$\psi = V\left\{1 + 2\left(\frac{V}{100}\right)^2\right\}^{\frac{1}{2}}$$

tabled for values of V from 0 to 50. The probable error of V must therefore be found by taking out the value of χ_2 corresponding to the given value of n, and multiplying it by the ψ found for the given value of V by interpolation from this table.

Thus suppose $V = 23{\cdot}65$ for 583 observations. We have by the usual advancing difference formula (*Biometrika*, Vol. II. p. 175)

$$\psi = 24{\cdot}18612 + {\cdot}65 \times 1{\cdot}15861 - \frac{{\cdot}65 \times {\cdot}35}{2} \times {\cdot}01316 + \frac{{\cdot}65 \times {\cdot}35 \times 1{\cdot}35}{2 \times 3} \times {\cdot}00045$$

which evaluated by the Brunsviga $= 24{\cdot}93774$*. Further, from Table II., $\chi_2 = \cdot01975$. Thus the probable error $= \psi\chi_2 = \cdot49252$. The value actually found by direct calculation is ·49259, the difference, which is of no importance for practical statistics, depending upon cutting off χ_2 at the fifth figure.

While Tables I. and II. are the work of the author, Table III. is due to Dr Raymond Pearl and J. Blakeman. All the values were found to seven figures, but it seemed sufficient for practical work to register them to five.

* Actual value 24·937,739, so that the ψ table is amply sufficient.

TABLES I. AND II. *Probable Errors of Means and Standard Deviations.*

n	χ_1	χ_2
1	·67449	·47694
2	·47694	·33724
3	·38942	·27536
4	·33724	·23847
5	·30164	·21329
6	·27536	·19471
7	·25493	·18026
8	·23847	·16862
9	·22483	·15898
10	·21329	·15082
11	·20337	·14380
12	·19471	·13768
13	·18707	·13228
14	·18026	·12747
15	·17415	·12314
16	·16862	·11923
17	·16359	·11567
18	·15898	·11241
19	·15474	·10942
20	·15082	·10665
21	·14719	·10408
22	·14380	·10168
23	·14064	·09945
24	·13768	·09735
25	·13490	·09539
26	·13228	·09353
27	·12981	·09179
28	·12747	·09013
29	·12525	·08856
30	·12314	·08708
31	·12114	·08566
32	·11923	·08431
33	·11741	·08302
34	·11567	·08179
35	·11401	·08062
36	·11241	·07949
37	·11088	·07841
38	·10942	·07737
39	·10800	·07637
40	·10665	·07541
41	·10534	·07448
42	·10408	·07359
43	·10286	·07273
44	·10168	·07190
45	·10055	·07110
46	·09945	·07032
47	·09838	·06957
48	·09735	·06884
49	·09636	·06813
50	·09539	·06745

n	χ_1	χ_2
51	·09445	·06678
52	·09353	·06614
53	·09265	·06551
54	·09179	·06490
55	·09095	·06431
56	·09013	·06373
57	·08934	·06317
58	·08856	·06262
59	·08781	·06209
60	·08708	·06157
61	·08636	·06107
62	·08566	·06057
63	·08498	·06009
64	·08431	·05962
65	·08366	·05916
66	·08302	·05871
67	·08240	·05827
68	·08179	·05784
69	·08120	·05742
70	·08062	·05700
71	·08005	·05660
72	·07949	·05621
73	·07894	·05582
74	·07841	·05544
75	·07788	·05507
76	·07737	·05471
77	·07687	·05435
78	·07637	·05400
79	·07589	·05366
80	·07541	·05332
81	·07494	·05299
82	·07448	·05267
83	·07403	·05235
84	·07359	·05204
85	·07316	·05173
86	·07273	·05143
87	·07231	·05113
88	·07190	·05084
89	·07150	·05056
90	·07110	·05027
91	·07071	·05000
92	·07032	·04972
93	·06994	·04946
94	·06957	·04919
95	·06920	·04893
96	·06884	·04868
97	·06848	·04843
98	·06813	·04818
99	·06779	·04793
100	·06745	·04769

n	χ_1	χ_2
101	·06711	·04746
102	·06678	·04722
103	·06646	·04699
104	·06614	·04677
105	·06582	·04654
106	·06551	·04632
107	·06521	·04611
108	·06490	·04589
109	·06460	·04568
110	·06431	·04547
111	·06402	·04527
112	·06373	·04507
113	·06345	·04487
114	·06317	·04467
115	·06290	·04447
116	·06262	·04428
117	·06236	·04409
118	·06209	·04391
119	·06183	·04372
120	·06157	·04354
121	·06132	·04336
122	·06107	·04318
123	·06082	·04300
124	·06057	·04283
125	·06033	·04266
126	·06009	·04249
127	·05985	·04232
128	·05962	·04216
129	·05939	·04199
130	·05916	·04183
131	·05893	·04167
132	·05871	·04151
133	·05849	·04136
134	·05827	·04120
135	·05805	·04105
136	·05784	·04090
137	·05763	·04075
138	·05742	·04060
139	·05721	·04045
140	·05700	·04031
141	·05680	·04017
142	·05660	·04002
143	·05640	·03988
144	·05621	·03974
145	·05601	·03961
146	·05582	·03947
147	·05563	·03934
148	·05544	·03920
149	·05526	·03907
150	·05507	·03894

TABLES I. AND II.—(*continued*).

n	χ_1	χ_2
151	·05489	·03881
152	·05471	·03868
153	·05453	·03856
154	·05435	·03843
155	·05418	·03831
156	·05400	·03819
157	·05383	·03806
158	·05366	·03794
159	·05349	·03782
160	·05332	·03771
161	·05316	·03759
162	·05299	·03747
163	·05283	·03736
164	·05267	·03724
165	·05251	·03713
166	·05235	·03702
167	·05219	·03691
168	·05204	·03680
169	·05188	·03669
170	·05173	·03658
171	·05158	·03647
172	·05143	·03637
173	·05128	·03626
174	·05113	·03616
175	·05099	·03605
176	·05084	·03595
177	·05070	·03585
178	·05056	·03575
179	·05041	·03565
180	·05027	·03555
181	·05013	·03545
182	·05000	·03535
183	·04986	·03526
184	·04972	·03516
185	·04959	·03507
186	·04946	·03497
187	·04932	·03488
188	·04919	·03478
189	·04906	·03469
190	·04893	·03460
191	·04880	·03451
192	·04868	·03442
193	·04855	·03433
194	·04843	·03424
195	·04830	·03415
196	·04818	·03407
197	·04806	·03398
198	·04793	·03389
199	·04781	·03381
200	·04769	·03372

n	χ_1	χ_2
201	·04757	·03364
202	·04746	·03356
203	·04734	·03347
204	·04722	·03339
205	·04711	·03331
206	·04699	·03323
207	·04688	·03315
208	·04677	·03307
209	·04666	·03299
210	·04654	·03291
211	·04643	·03283
212	·04632	·03276
213	·04622	·03268
214	·04611	·03260
215	·04600	·03253
216	·04589	·03245
217	·04579	·03238
218	·04568	·03230
219	·04558	·03223
220	·04547	·03216
221	·04537	·03208
222	·04527	·03201
223	·04517	·03194
224	·04507	·03187
225	·04497	·03180
226	·04487	·03173
227	·04477	·03166
228	·04467	·03159
229	·04457	·03152
230	·04447	·03145
231	·04438	·03138
232	·04428	·03131
233	·04419	·03125
234	·04409	·03118
235	·04400	·03111
236	·04391	·03105
237	·04381	·03098
238	·04372	·03092
239	·04363	·03085
240	·04354	·03079
241	·04345	·03172
242	·04336	·03066
243	·04327	·03060
244	·04318	·03053
245	·04309	·03047
246	·04300	·03041
247	·04292	·03035
248	·04283	·03029
249	·04274	·03022
250	·04266	·03016

n	χ_1	χ_2
251	·04257	·03010
252	·04249	·03004
253	·04240	·02998
254	·04232	·02993
255	·04224	·02987
256	·04216	·02981
257	·04207	·02975
258	·04199	·02969
259	·04191	·02964
260	·04183	·02958
261	·04175	·02952
262	·04167	·02947
263	·04159	·02941
264	·04151	·02935
265	·04143	·02930
266	·04136	·02924
267	·04128	·02919
268	·04120	·02913
269	·04112	·02908
270	·04105	·02903
271	·04097	·02897
272	·04090	·02892
273	·04082	·02887
274	·04075	·02881
275	·04067	·02876
276	·04060	·02871
277	·04053	·02866
278	·04045	·02860
279	·04038	·02855
280	·04031	·02850
281	·04024	·02845
282	·04017	·02840
283	·04009	·02835
284	·04002	·02830
285	·03995	·02825
286	·03988	·02820
287	·03981	·02815
288	·03974	·02810
289	·03968	·02806
290	·03961	·02801
291	·03954	·02796
292	·03947	·02791
293	·03940	·02786
294	·03934	·02782
295	·03927	·02777
296	·03920	·02772
297	·03913	·02767
298	·03907	·02763
299	·03901	·02758
300	·03894	·02754

TABLES I. AND II.—(*continued*).

n	X_1	X_2
301	·03888	·02749
302	·03881	·02744
303	·03875	·02740
304	·03868	·02735
305	·03862	·02731
306	·03856	·02726
307	·03850	·02722
308	·03843	·02718
309	·03837	·02713
310	·03831	·02709
311	·03825	·02704
312	·03819	·02700
313	·03812	·02696
314	·03806	·02692
315	·03800	·02687
316	·03794	·02683
317	·03788	·02679
318	·03782	·02675
319	·03776	·02670
320	·03771	·02666
321	·03765	·02662
322	·03759	·02658
323	·03753	·02654
324	·03747	·02650
325	·03741	·02646
326	·03736	·02642
327	·03730	·02637
328	·03724	·02633
329	·03719	·02629
330	·03713	·02625
331	·03707	·02621
332	·03702	·02618
333	·03696	·02614
334	·03691	·02610
335	·03685	·02606
336	·03680	·02602
337	·03674	·02598
338	·03669	·02594
339	·03663	·02590
340	·03658	·02587
341	·03653	·02583
342	·03647	·02579
343	·03642	·02575
344	·03637	·02571
345	·03631	·02568
346	·03626	·02564
347	·03621	·02560
348	·03616	·02557
349	·03610	·02553
350	·03605	·02549

n	X_1	X_2
351	·03600	·02546
352	·03595	·02542
353	·03590	·02538
354	·03585	·02535
355	·03580	·02531
356	·03575	·02528
357	·03570	·02524
358	·03565	·02521
359	·03560	·02517
360	·03555	·02514
361	·03550	·02510
362	·03545	·02507
363	·03540	·02503
364	·03535	·02500
365	·03530	·02496
366	·03526	·02493
367	·03521	·02490
368	·03516	·02486
369	·03511	·02483
370	·03507	·02479
371	·03502	·02476
372	·03497	·02473
373	·03492	·02469
374	·03488	·02466
375	·03483	·02463
376	·03478	·02460
377	·03474	·02456
378	·03469	·02453
379	·03465	·02450
380	·03460	·02447
381	·03456	·02443
382	·03451	·02440
383	·03446	·02437
384	·03442	·02434
385	·03438	·02431
386	·03433	·02428
387	·03429	·02424
388	·03424	·02421
389	·03420	·02418
390	·03415	·02415
391	·03411	·02412
392	·03407	·02409
393	·03402	·02406
394	·03398	·02403
395	·03394	·02400
396	·03389	·02397
397	·03385	·02394
398	·03381	·02391
399	·03377	·02388
400	·03372	·02385

n	X_1	X_2
401	·03368	·02382
402	·03364	·02379
403	·03360	·02376
404	·03356	·02373
405	·03352	·02370
406	·03347	·02367
407	·03343	·02364
408	·03339	·02361
409	·03335	·02358
410	·03331	·02355
411	·03327	·02353
412	·03323	·02350
413	·03319	·02347
414	·03315	·02344
415	·03311	·02341
416	·03307	·02338
417	·03303	·02336
418	·03299	·02333
419	·03295	·02330
420	·03291	·02327
421	·03287	·02324
422	·03283	·02322
423	·03279	·02319
424	·03276	·02316
425	·03272	·02313
426	·03268	·02311
427	·03264	·02308
428	·03260	·02305
429	·03256	·02303
430	·03253	·02300
431	·03249	·02297
432	·03245	·02295
433	·03241	·02292
434	·03238	·02289
435	·03234	·02287
436	·03230	·02284
437	·03227	·02281
438	·03223	·02279
439	·03219	·02276
440	·03216	·02274
441	·03212	·02271
442	·03208	·02269
443	·03205	·02266
444	·03201	·02263
445	·03197	·02261
446	·03194	·02258
447	·03190	·02256
448	·03187	·02253
449	·03183	·02251
450	·03180	·02248

TABLES I. AND II.—(*continued*).

n	χ_1	χ_2
451	·03176	·02246
452	·03173	·02243
453	·03169	·02241
454	·03166	·02238
455	·03162	·02236
456	·03159	·02233
457	·03155	·02231
458	·03152	·02229
459	·03148	·02226
460	·03145	·02224
461	·03141	·02221
462	·03138	·02219
463	·03135	·02217
464	·03131	·02214
465	·03128	·02212
466	·03125	·02209
467	·03121	·02207
468	·03118	·02205
469	·03115	·02202
470	·03111	·02200
471	·03108	·02198
472	·03105	·02195
473	·03101	·02193
474	·03098	·02191
475	·03095	·02188
476	·03092	·02186
477	·03088	·02184
478	·03085	·02181
479	·03082	·02179
480	·03079	·02177
481	·03075	·02175
482	·03072	·02172
483	·03069	·02170
484	·03066	·02168
485	·03063	·02166
486	·03060	·02163
487	·03056	·02161
488	·03053	·02159
489	·03050	·02157
490	·03047	·02155
491	·03044	·02152
492	·03041	·02150
493	·03038	·02148
494	·03035	·02146
495	·03032	·02144
496	·03029	·02142
497	·03026	·02139
498	·03022	·02137
499	·03019	·02135
500	·03016	·02133

n	χ_1	χ_2
501	·03013	·02131
502	·03010	·02129
503	·03007	·02127
504	·03004	·02124
505	·03001	·02122
506	·02998	·02120
507	·02996	·02118
508	·02993	·02116
509	·02990	·02114
510	·02987	·02112
511	·02984	·02110
512	·02981	·02108
513	·02978	·02106
514	·02975	·02104
515	·02972	·02102
516	·02969	·02100
517	·02966	·02098
518	·02964	·02096
519	·02961	·02094
520	·02958	·02092
521	·02955	·02089
522	·02952	·02087
523	·02949	·02085
524	·02947	·02084
525	·02944	·02082
526	·02941	·02080
527	·02938	·02078
528	·02935	·02076
529	·02933	·02074
530	·02930	·02072
531	·02927	·02070
532	·02924	·02068
533	·02922	·02066
534	·02919	·02064
535	·02916	·02062
536	·02913	·02060
537	·02911	·02058
538	·02908	·02056
539	·02905	·02054
540	·02903	·02052
541	·02900	·02051
542	·02897	·02049
543	·02895	·02047
544	·02892	·02045
545	·02889	·02043
546	·02887	·02041
547	·02884	·02039
548	·02881	·02037
549	·02879	·02036
550	·02876	·02034

n	χ_1	χ_2
551	·02873	·02032
552	·02871	·02030
553	·02868	·02028
554	·02866	·02026
555	·02863	·02024
556	·02860	·02023
557	·02858	·02021
558	·02855	·02019
559	·02853	·02017
560	·02850	·02015
561	·02848	·02014
562	·02845	·02012
563	·02843	·02010
564	·02840	·02008
565	·02838	·02006
566	·02835	·02005
567	·02833	·02003
568	·02830	·02001
569	·02828	·01999
570	·02825	·01998
571	·02823	·01996
572	·02820	·01994
573	·02818	·01992
574	·02815	·01991
575	·02813	·01990
576	·02810	·01987
577	·02808	·01986
578	·02806	·01984
579	·02803	·01982
580	·02801	·01980
581	·02798	·01978
582	·02796	·01977
583	·02793	·01975
584	·02791	·01974
585	·02789	·01972
586	·02786	·01970
587	·02784	·01969
588	·02782	·01967
589	·02779	·01965
590	·02777	·01964
591	·02774	·01962
592	·02772	·01960
593	·02770	·01959
594	·02767	·01957
595	·02765	·01955
596	·02763	·01954
597	·02761	·01952
598	·02758	·01950
599	·02756	·01949
600	·02754	·01947

TABLES I. AND II.—(*continued*).

n	X_1	X_2
601	·02751	·01945
602	·02749	·01944
603	·02747	·01942
604	·02744	·01941
605	·02742	·01939
606	·02740	·01937
607	·02738	·01936
608	·02735	·01934
609	·02733	·01933
610	·02731	·01931
611	·02729	·01929
612	·02726	·01928
613	·02724	·01926
614	·02722	·01925
615	·02720	·01923
616	·02718	·01922
617	·02715	·01920
618	·02713	·01919
619	·02711	·01917
620	·02709	·01915
621	·02707	·01914
622	·02704	·01912
623	·02702	·01911
624	·02700	·01909
625	·02698	·01908
626	·02696	·01906
627	·02694	·01905
628	·02692	·01903
629	·02689	·01902
630	·02687	·01900
631	·02685	·01899
632	·02683	·01897
633	·02681	·01896
634	·02679	·01894
635	·02677	·01893
636	·02675	·01891
637	·02672	·01890
638	·02670	·01888
639	·02668	·01887
640	·02666	·01885
641	·02664	·01884
642	·02662	·01822
643	·02660	·01881
644	·02658	·01879
645	·02656	·01878
646	·02654	·01876
647	·02652	·01875
648	·02650	·01874
649	·02648	·01872
650	·02646	·01871

n	X_1	X_2
651	·02644	·01869
652	·02642	·01868
653	·02639	·01866
654	·02637	·01865
655	·02635	·01864
656	·02633	·01862
657	·02631	·01861
658	·02629	·01859
659	·02627	·01858
660	·02625	·01856
661	·02623	·01855
662	·02621	·01854
663	·02620	·01852
664	·02618	·01851
665	·02616	·01849
666	·02614	·01848
667	·02612	·01847
668	·02610	·01845
669	·02608	·01844
670	·02606	·01843
671	·02604	·01841
672	·02602	·01840
673	·02600	·01838
674	·02598	·01837
675	·02596	·01836
676	·02594	·01834
677	·02592	·01833
678	·02590	·01832
679	·02588	·01830
680	·02587	·01829
681	·02585	·01828
682	·02583	·01826
683	·02581	·01825
684	·02579	·01824
685	·02577	·01822
686	·02575	·01821
687	·02573	·01820
688	·02571	·01818
689	·02570	·01817
690	·02568	·01816
691	·02566	·01814
692	·02564	·01813
693	·02562	·01812
694	·02560	·01810
695	·02558	·01809
696	·02557	·01808
697	·02555	·01807
698	·02553	·01805
699	·02551	·01804
700	·02549	·01803

n	X_1	X_2
701	·02548	·01801
702	·02546	·01800
703	·02544	·01799
704	·02542	·01798
705	·02540	·01796
706	·02538	·01795
707	·02537	·01794
708	·02535	·01792
709	·02533	·01791
710	·02531	·01790
711	·02530	·01789
712	·02528	·01787
713	·02526	·01786
714	·02524	·01785
715	·02522	·01784
716	·02521	·01782
717	·02519	·01781
718	·02517	·01780
719	·02515	·01779
720	·02514	·01777
721	·02512	·01776
722	·02510	·01775
723	·02508	·01774
724	·02507	·01773
725	·02505	·01771
726	·02503	·01770
727	·02502	·01769
728	·02500	·01768
729	·02498	·01766
730	·02496	·01765
731	·02495	·01764
732	·02493	·01763
733	·02491	·01762
734	·02490	·01760
735	·02488	·01759
736	·02486	·01758
737	·02485	·01757
738	·02483	·01756
739	·02481	·01754
740	·02479	·01753
741	·02478	·01752
742	·02476	·01751
743	·02474	·01750
744	·02473	·01749
745	·02471	·01747
746	·02469	·01746
747	·02468	·01745
748	·02466	·01744
749	·02465	·01743
750	·02463	·01742

TABLES I. AND II.—(*continued*).

n	X_1	X_2
751	·02461	·01740
752	·02460	·01739
753	·02458	·01738
754	·02456	·01737
755	·02455	·01736
756	·02453	·01735
757	·02451	·01733
758	·02450	·01732
759	·02448	·01731
760	·02447	·01730
761	·02445	·01729
762	·02443	·01728
763	·02442	·01727
764	·02440	·01725
765	·02439	·01724
766	·02437	·01723
767	·02435	·01722
768	·02434	·01721
769	·02432	·01720
770	·02431	·01719
771	·02429	·01718
772	·02428	·01717
773	·02426	·01715
774	·02424	·01714
775	·02423	·01713
776	·02421	·01712
777	·02420	·01711
778	·02418	·01710
779	·02417	·01709
780	·02415	·01708
781	·02414	·01707
782	·02412	·01706
783	·02410	·01704
784	·02409	·01703
785	·02407	·01702
786	·02406	·01701
787	·02404	·01700
788	·02403	·01699
789	·02401	·01698
790	·02400	·01697
791	·02398	·01696
792	·02397	·01695
793	·02395	·01694
794	·02394	·01693
795	·02392	·01692
796	·02391	·01690
797	·02389	·01689
798	·02388	·01688
799	·02386	·01687
800	·02385	·01686

n	X_1	X_2
801	·02383	·01685
802	·02382	·01684
803	·02380	·01683
804	·02379	·01682
805	·02377	·01681
806	·02376	·01680
807	·02374	·01679
808	·02373	·01678
809	·02371	·01677
810	·02370	·01676
811	·02368	·01675
812	·02367	·01674
813	·02366	·01673
814	·02364	·01672
815	·02363	·01671
816	·02361	·01670
817	·02360	·01669
818	·02358	·01668
819	·02357	·01667
820	·02355	·01666
821	·02354	·01665
822	·02353	·01664
823	·02351	·01662
824	·02350	·01661
825	·02348	·01660
826	·02347	·01659
827	·02345	·01658
828	·02344	·01657
829	·02343	·01656
830	·02341	·01655
831	·02340	·01654
832	·02338	·01653
833	·02337	·01652
834	·02336	·01651
835	·02334	·01651
836	·02333	·01650
837	·02331	·01649
838	·02330	·01648
839	·02329	·01647
840	·02327	·01646
841	·02326	·01645
842	·02324	·01644
843	·02323	·01643
844	·02322	·01642
845	·02320	·01641
846	·02319	·01640
847	·02318	·01639
848	·02316	·01638
849	·02315	·01637
850	·02313	·01636

n	X_1	X_2
851	·02312	·01635
852	·02311	·01634
853	·02309	·01633
854	·02308	·01632
855	·02307	·01631
856	·02305	·01630
857	·02304	·01629
858	·02303	·01628
859	·02301	·01627
860	·02300	·01626
861	·02299	·01625
862	·02297	·01624
863	·02296	·01624
864	·02295	·01623
865	·02293	·01622
866	·02292	·01621
867	·02291	·01620
868	·02289.	·01619
869	·02288	·01618
870	·02287	·01617
871	·02285	·01616
872	·02284	·01615
873	·02283	·01614
874	·02281	·01613
875	·02280	·01612
876	·02279	·01611
877	·02278	·01610
878	·02276	·01610
879	·02275	·01609
880	·02274	·01608
881	·02272	·01607
882	·02271	·01606
883	·02270	·01605
884	·02269	·01604
885	·02267	·01603
886	·02266	·01602
887	·02265	·01601
888	·02263	·01600
889	·02262	·01600
890	·02261	·01599
891	·02260	·01598
892	·02258	·01597
893	·02257	·01596
894	·02256	·01595
895	·02255	·01594
896	·02253	·01593
897	·02252	·01592
898	·02251	·01592
899	·02250	·01591
900	·02248	·01590

TABLES I. AND II.—(*continued*).

n	χ_1	χ_2	n	χ_1	χ_2
901	·02247	·01589	951	·02187	·01547
902	·02246	·01588	952	·02186	·01546
903	·02245	·01587	953	·02185	·01545
904	·02243	·01586	954	·02184	·01544
905	·02242	·01585	955	·02183	·01543
906	·02241	·01585	956	·02181	·01543
907	·02240	·01584	957	·02180	·01542
908	·02238	·01583	958	·02179	·01541
909	·02237	·01582	959	·02178	·01540
910	·02236	·01581	960	·02177	·01539
911	·02235	·01580	961	·02176	·01539
912	·02233	·01579	962	·02175	·01538
913	·02232	·01578	963	·02174	·01537
914	·02231	·01578	964	·02172	·01536
915	·02230	·01577	965	·02171	·01535
916	·02229	·01576	966	·02170	·01535
917	·02227	·01575	967	·02169	·01534
918	·02226	·01574	968	·02168	·01533
919	·02225	·01573	969	·02167	·01532
920	·02224	·01572	970	·02166	·01531
921	·02223	·01572	971	·02165	·01531
922	·02221	·01571	972	·02163	·01530
923	·02220	·01570	973	·02162	·01529
924	·02219	·01569	974	·02161	·01528
925	·02218	·01568	975	·02160	·01527
926	·02217	·01567	976	·02159	·01527
927	·02215	·01566	977	·02158	·01526
928	·02214	·01566	978	·02157	·01525
929	·02213	·01565	979	·02156	·01524
930	·02212	·01564	980	·02155	·01524
931	·02211	·01563	981	·02153	·01523
932	·02209	·01562	982	·02152	·01522
933	·02208	·01561	983	·02151	·01521
934	·02207	·01561	984	·02150	·01520
935	·02206	·01560	985	·02149	·01520
936	·02205	·01559	986	·02148	·01519
937	·02203	·01558	987	·02147	·01518
938	·02202	·01557	988	·02146	·01517
939	·02201	·01556	989	·02145	·01517
940	·02200	·01556	990	·02144	·01516
941	·02199	·01555	991	·02143	·01515
942	·02198	·01554	992	·02142	·01514
943	·02196	·01553	993	·02140	·01514
944	·02195	·01552	994	·02139	·01513
945	·02194	·01551	995	·02138	·01512
946	·02193	·01551	996	·02137	·01511
947	·02192	·01550	997	·02136	·01510
948	·02191	·01549	998	·02135	·01510
949	·02189	·01548	999	·02134	·01509
950	·02188	·01547	1000	·02133	·01508

TABLE III.

Probable Errors of Coefficients of Variation.

V	ψ	Δ +	Δ^2 +	Δ^3 +
0	0·00000	1·00010	0	60
1	1·00010	1·00070	60	60
2	2·00080	1·00190	120	60
3	3·00270	1·00370	180	60
4	4·00639	1·00609	239	59
5	5·01248	1·00908	299	59
6	6·02156	1·01266	358	59
7	7·03422	1·01682	417	58
8	8·05104	1·02157	475	58
9	9·07261	1·02690	533	57
10	10·09950	1·03280	590	57
11	11·13230	1·03927	647	56
12	12·17157	1·04630	703	56
13	13·21787	1·05389	759	55
14	14·27176	1·06202	814	54
15	15·33379	1·07070	868	53
16	16·40449	1·07991	921	53
17	17·48440	1·08965	974	52
18	18·57405	1·09990	1025	51
19	19·67395	1·11066	1076	50
20	20·78461	1·12192	1126	49
21	21·90653	1·13368	1175	48
22	23·04021	1·14591	1223	47
23	24·18612	1·15861	1270	46
24	25·34473	1·17177	1316	45
25	26·51650	1·18539	1362	44
26	27·70190	1·19945	1406	43
27	28·90135	1·21395	1449	42
28	30·11530	1·22886	1491	41
29	31·34416	1·24418	1533	40
30	32·58834	1·25991	1573	39
31	33·84825	1·27603	1612	38
32	35·12428	1·29253	1650	37
33	36·41681	1·30940	1687	36
34	37·72621	1·32664	1723	35
35	39·05285	1·34422	1758	34
36	40·39707	1·36215	1793	33
37	41·75922	1·38041	1826	32
38	43·13962	1·39899	1858	31
39	44·53861	1·41789	1890	30
40	45·95650	1·43709	1920	30
41	47·39359	1·45658	1950	29
42	48·85017	1·47636	1978	28
43	50·32654	1·49642	2006	27
44	51·82296	1·51675	2033	26
45	53·33971	1·53734	2059	25
46	54·87706	1·55818	2084	24
47	56·43524	1·57927	2109	24
48	58·01451	1·60059	2132	23
49	59·61510	1·62214	2155	22
50	61·23724		2177	

COOPERATIVE INVESTIGATIONS ON PLANTS.

III. ON INHERITANCE IN THE SHIRLEY POPPY*.

SECOND MEMOIR.

CONTENTS.

		Page
(1)	Material and Manner of Fertilisation	394
(2)	Nature of the Observations	398
(3)	Influence of Environment on the Shirley Characters	401
(4)	Influence of Selection of Maternal Plants	403
(5)	Influence of Variability of like Organs—Homotyposis—On Coefficients of Heredity	409
(6)	Parental Correlation	410
(7)	Fraternal Correlation	413
(8)	General Conclusions	415
	Appendix: Correlation Tables	416

(1) *Material and Manner of Fertilisation.*

A first memoir on the subject of inheritance in the Shirley Poppy was published in November, 1902†. In summing up our results in that paper we stated that:

"Most plant organs being multiple in appearance, we have to apply special methods to deduce the intensity of heredity from multiple observations on the individual. It will probably be better in future experiments to confine attention to the first or principal flower, instead of using the indirect method of homotyposis; but this will involve the observation (previous to harvesting) of individual plants in large series—500 to 1000—and much increase the labour of superintendence and observation‡."

We continued:

"Notwithstanding the difficulties referred to above we find that for a variety of plant characters in the Shirley Poppy the values of hereditary influence found are on the whole in fair agreement with the like values in man."

* These experiments have been materially assisted by the Grant made to the Department of Applied Mathematics at University College, London, by the Worshipful Company of Drapers. The establishment of a paid calculating staff has in addition largely accelerated the labour of reduction.

† *Biometrika*, Vol. II. pp. 56–100. ‡ *Ibid.*, p. 83.

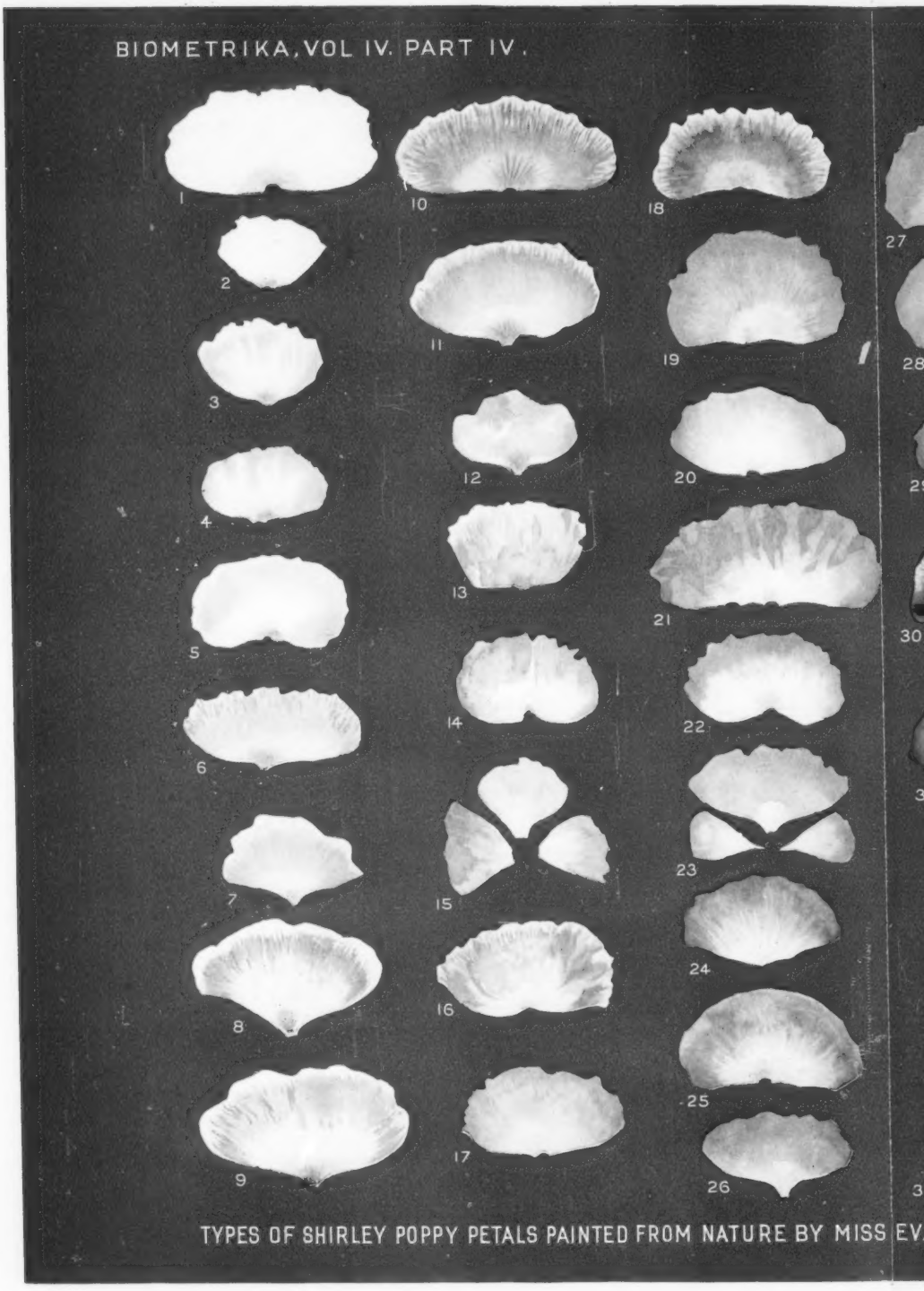
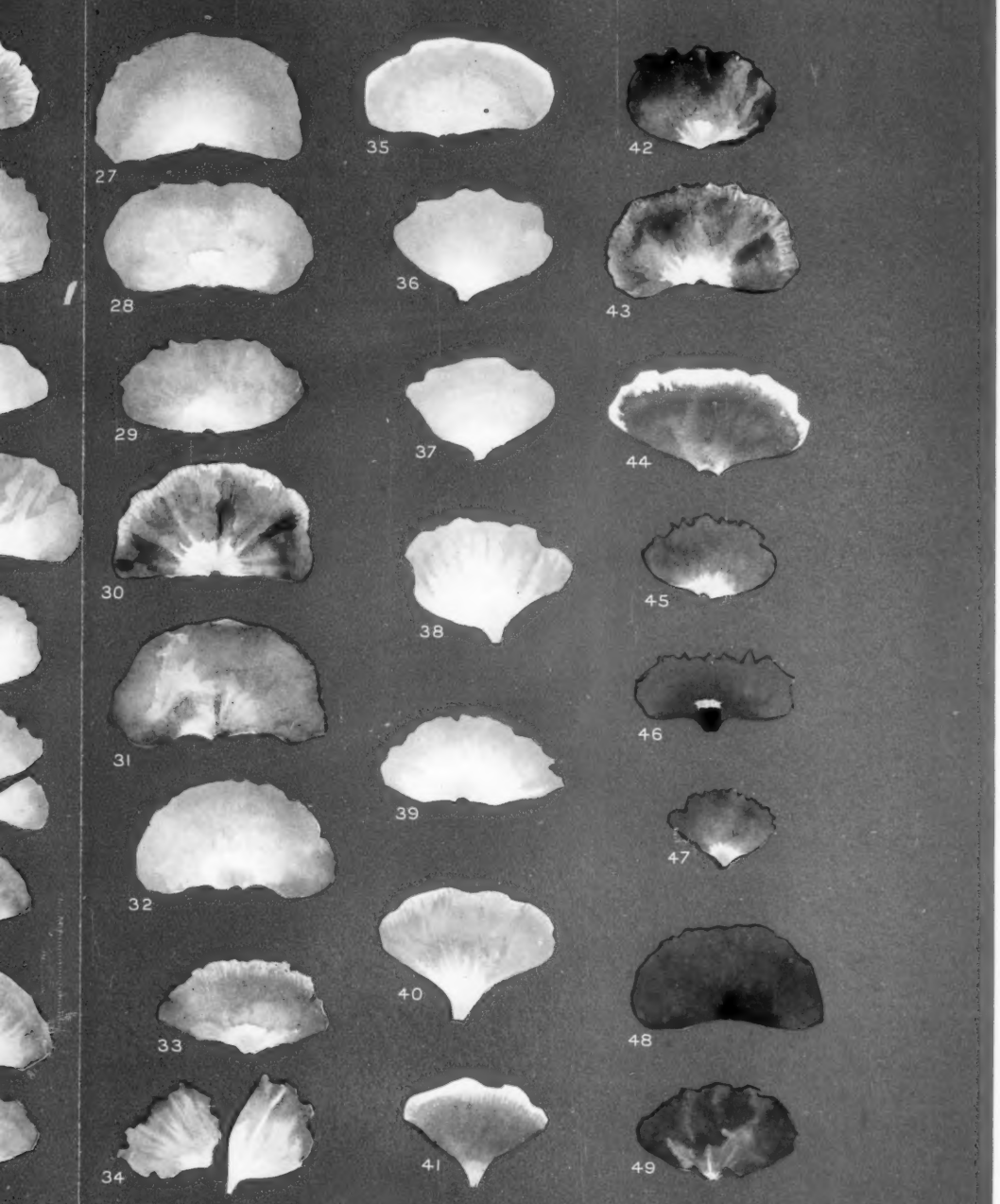

BIOMETRIKA, VOL IV. PART IV.
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
30
TYPES OF SHIRLEY POPPY PETALS PAINTED FROM NATURE BY MISS EV

PLATE I.



Y MISS EVA BIGGS FOR THE MEMOIR ON INHERITANCE IN THE SHIRLEY POPPY.

The object of the first series of experiments, like those which have followed it, was to ascertain the intensity of heredity in a free plant *population*, and to compare such heredity with that of similar populations in other types of life. If we have to deal with evolution in actual living forms in which isolation of individuals is not the rule, it is the law of inheritance in *populations* with which we are concerned. We desire to know the statistics of sub-classes, and the proportions in which those sub-classes reproduce themselves. A statistical investigation of this kind cannot be opposed to any ultimate analysis of the result of crossing *individuals*; it can only indeed serve as a criterion of the truth of any hypothesis based upon individual crossings. At the same time it avoids the dangers which arise from neglecting what we may term the "law of small numbers." It is too often forgotten that an imposing display of numbers, obtained by dealing with the numerous offspring of a number of crossings which can be counted on the fingers, must on any hypothesis of parental gametic constancy be given the weight of the number of crossings and not of the number of offspring, when the results are applied to the race as a whole.

In our earlier experiments we were not certain to what extent Shirley Poppies grown in masses were cross or self-fertilised *. This doubt left a double interpretation open when we came to deal with the resemblance of siblings. Our results in the previous investigation were in general accordance with those for man, provided we assumed that the bulk of the plants were *cross-fertilised*.

Another point of considerable importance was the question of whether any really well-defined colour sub-groups could be asserted to exist in the case of the Shirley Poppy. In order to determine this point water-colour sketches of the petals of a long series of the poppies were kindly made for us by Miss Eva Biggs immediately after gathering. It was soon seen that in distribution of colour, amount of colour, and shade of pigmentation there was practical continuity. We found it impossible to draw a rigid line between scarlets, reds, orange-reds, terra-cottas, rose-pinks, pinks, and pink-whites, although such terminology served fairly well for rough classification. There appeared always to be transitional individuals, although such individuals might be numerically few. The ocular demonstration of this was difficult even in Miss Biggs' paintings, because the colour of the poppy is very difficult to match in water-colours, and being matched is very evanescent. We believe that Miss Biggs' careful work fully achieves what is possible in this direction. But the task becomes harder still when the colour printer has to match the artist's tints. If honest, he frankly tells you that he can approximately catch the artist's shades, but that he can only do so by using non-durable printer's inks. The alternative before the Editors was therefore to produce something like the artist's work which will not be permanent, or something unlike it which will be permanent. The latter seems to fail entirely in its object, for while it might show continuity, it would give colours which we should be told were not even approximately those of the Shirley. We have therefore determined on the second course,

* *Loc. cit.*, p. 62.

namely, to allow the colour-printer to reproduce in possibly evanescent colours, which will show for some time at least the apparent continuity of pigmentation in the Shirley Poppy. It will be seen from our plate that while white-pinks, pinks, rose-pinks, terra-cottas*, and reds are fairly safe broad classifications, there is no absolute discontinuity when we come to deal with individual instances.

In the present experiments the observations were all made on the first or apical flower. They had to be made while the plant was in blossom. The seed of the apical capsule observed was taken and marked with the new pedigree number. A second and sometimes a third bud on the same plant was enclosed in a bag of very fine bolting cloth, or for about half the experiments in oiled paper bags obtained from Cologne. These were invariably used in the second year's experiments. The advantages of the bolting cloth bag are the free admission of air and light, but it becomes heavy in rainy weather and much damage was done in the wet season of 1903. It is possible that very minute insects may occasionally penetrate it, but the risk here is not much greater than with paper bags, however carefully they are tied round the stem. In the spring of 1903 series of 100 packets of pedigree seed were issued to a number of cooperators. Each of these packets contained seed from one *plant* only, and the ultimate origin of the seed was the 1899 Hampden crop of Pearson, although there had been several intervening generations. In the spring of 1904 series of 100 packets of apical capsule seeds of 1903 and of 50 packets of bagged capsule seeds of 1903 were issued, and the same process was followed of collecting the seed of the apical capsules and of a secondary bagged capsule. To test fertility the seed in the free and bagged capsules was merely examined after gathering and not again sown.

The following are the 1903 *crops:*

(A) *Parkstone Crop.* This was grown and observed by H. Blanchard. The crop was sown on May 13th, the first seedlings were observed on the 20th, and the thinning took place between June 13 and 15. Of the 100 packets of single capsule seed provided all produced plants and in most cases more than sufficient plants to get 10 siblings. In one case only one seed and in a second three seeds alone germinated†. But the observer was able to deal with nearly 1000 individual plants. The weather was extremely unfavourable during the harvesting and much harm was done to the bagged capsules. Tabulation and calculations by Alice Lee, D.Sc.

(B) *Arncliffe Crop.* This was grown and observed by the Rev. W. A. Shuffry, and 50 packets of seed were sown on May 30; first seedlings observed June 10, thinned July 13 and 20. Second series of 50 sown June 18, first seed-

* In the original Hampden Shirley Poppies the terra-cotta was extremely rare, if it occurred at all. It is still scarce, but is creeping up in numbers, and may ultimately need a special class. As meanwhile it is being introduced commercially, we cannot say whether it is a natural evolution or has been introduced by cross-fertilisation into our crops.

† One of these plants also gave a distorted apical capsule with a "stigmatic labyrinth" instead of bands.

lings observed June 25, thinned July 29–30. Out of the earlier sowing, none; out of the later sowing four sets failed to germinate. There were, however, a number of series in which less than ten plants failed to appear or be observed. The lateness of the Dale* season and the badness of the weather were much against this crop. Tabulation and calculation by J. Clerk-Maxwell Garnett.

(C) *Streatham Crop.* This was grown and observed by W. Palin Elderton. Sown on April 30, thinned at end of July. There was no case in which none of the seed of the 95 packets sown failed to germinate, but in some cases very few plants were grown. The weather was very wet during the whole of the flowering season and also very windy; many of the plants were much damaged. The bagged capsules suffered very much, and the seed had to be gathered earlier than was desirable. The tabulation is due to Mary Beeton, and the calculations to Alice Lee and her conjointly.

(D) *Boar's Hill Crop.* This was grown and observed by A. D. Darbishire. Sown April 23–25, first seedlings observed May 7, thinned June 25 to July 1. Out of 42 packets of seed sown at least 40 germinated. But unfortunately the observations were not made on the principal flower, but on the 2nd, 3rd, 4th, and even 5th flowers, so that the material is not comparable with the other observations of this series. Not yet tabled or calculated†.

(E) *Kidderminster Crop.* This was grown and observed by John Notcutt. Seed sown April 17, first seedlings observed April 29, thinned June 25. Of the 100 packets sown all germinated, and there were nearly 1000 plants observed. The tabulation and calculations are by Mary Beeton.

With regard to the seed provided for crops (A) to (E), it must be remarked that in making up the packets fertile capsules were directly chosen and those capsules which provided unfertilised seed deliberately rejected. This method could not be followed in the seed for the 1904 crops, because the harvest showed that the bagged capsules contained apparently no fertile seed. Accordingly the capsules for the seed-sowing in 1904 were taken from the 1903 (A) crop at random, each capsule having been at harvest placed in a bag whether or no it contained fertile seed.

The 1904 *crops were:*

(F) *Crewe Crop.* This was grown and observed by H. E. Ward. Sown May 18, first seedlings observed May 27, thinned June 30. Out of 129 packets of free apical capsule seed 92 germinated. This may be taken to represent the reduction in fertility due to no selection of fertile seed capsules. In 1903 there was practically no failure to germinate. Out of 50 bagged capsule packets 46 failed entirely to germinate, and 4 germinated, giving respectively 5, 2, 1, and 1

* Arncliffe is in Upper Wharfdale.

† It is hoped that enough material may ultimately be accumulated to test the difference between sibling resemblance measured on apical and on other flowers.

actually developed plants; the 92 which germinated of the free capsules gave 400 *actually observed* plants*. The tabulation and calculations are due to Dr A. Lee.

(G) *Bushey Crop.* Grown by H. Payne, but not recorded owing to his appointment to the South African College. The bagged capsules, however, failed to give seed which germinated, while the free capsules gave a very good crop.

(H) *Rushden Crop.* Grown by G. Abbott. Mr Abbott's time was so occupied by other matters during the flowering season that few records were made of the flowers, but the same result occurred in this sowing as at Bushey—the bagged capsules gave no germinating seed, the free capsules a plentiful supply.

We may therefore conclude that while a plentiful supply of fertile seed is obtained when free capsules growing in the midst of a poppy crop are collected, practically no seed is obtained when we deal with bagged capsules, at any rate in seasons like 1903 and 1904. It seems reasonable therefore to infer that the Shirley Poppy is in the bulk cross and not self-fertilised†, and we should accordingly expect in the resemblance of sister plants the intensity to approach what we are familiar with in other types of life as that of *half* and *not whole* siblings. We may remark that the observers noted continual visits of wasps, bees, and other insects to the crops. As far as could be determined from a careful examination of the unused bagged capsules of 1903, they presented no fertile seed whatever.

(2) *Nature of the Observations.*

A full crop consisted of 100 to 150 plots each sown with seed from a single pedigree individual plant. In each plot at a maximum ten and of necessity sometimes fewer individual plants were selected and the characters of their apical flowers observed. The characters were as follows:

(i) *Petals.* (*a*) The number of petals. (*b*) The number of petaloid stamens.

(ii) *Number of stigmatic bands.* These being counted at the periphery and not at the centre.

(iii) *Colour of the "middle third" of the petals.* This was recorded for one petal of the flower only, unless the petals were unlike, when the colours of both petal pairs were given. The colour categories used were:

Red (*r*).	Pink-white (*pw*).	White (*w*).
Red-pink (*rp*).	Veined-white (*vw*).	Red-white (*rw*).
Pink (*p*).		

* If these beds had not been thinned, and if all the plants in some of the beds had been observed, there would have been far more, i.e. 400 represents the minimum, 9 the maximum fertility of the two groups.

† Another, we consider fairly strong, argument in favour of wide cross-fertilisation can be found in the occurrence of poppies with dark pollen or dark basal patches. Such "rogues" occur where there is no trace of black in the maternal plant, but they generally appear as isolated units; sometimes, but rarely, two or three cases appear in a fraternity. If the whole fraternity were due to one pollen parent, they could hardly fail to be more frequent in the group.

Scarlet was tabled under red, the few terra-cottas and salmon-pink under red-pink, and intermediate colours by intermediate values (e.g. *p* or *pw*). If the petal pairs were of different colours, they were recorded as $p \times w$, say, and tabled under intermediate values. Veined-white was used for white faintly pink or for white with distinct pink veins.

(iv) *Breadth of margin.* The classes used were:

Broad (*b*).	Slight (*s*).	None (*n*).
Broad to slight (*bs*).	Slight to none (*sn*).	

"None" signifies that the border of the petal was of the same colour as its body or middle-third.

(v) *Colour of the margin, if it existed.* This was only completely noted in the Parkstone crop, and it enabled us to form a classification of the flowers by colour at the border of the petal; this was the same as in the middle-third, if there were no separate margin. Colour scale used as in (iii).

(vi) *Intensity of the wrinkling of the petal.* This was observed some time, 12 to 24 hours, after the opening of the flower. But while frilling is a marked feature of the Shirley, this character was not a satisfactory one to observe, and but little weight can be given to the results. The following classes were adopted:

Very much wrinkled or frilled (*f*).	Slightly wrinkled (*sw*).
Wrinkled (*w*).	Not wrinkled (*nw*).

(vii) *Base of petal.*

(α) If the petal base differ in colour from the middle-third it may be said to have a basal patch, and the following classification of basal patches was adopted:

None (*n*).	Well defined (*d*).
None to slight (*ns*).	Well defined to large (*dl*).
Slight (*s*).	Large (*l*).
Slight to well defined (*sd*).	

(β) If there were a basal patch its colour was noted; white, black and lilac basal patches were recorded.

(viii) The shape of the capsule cover, whether flat or conical, the colour of the pollen, and all abnormalities were further noted in a separate column. Actual shape of capsules was only fully noted in the Crewe crop.

The following are samples of entries from the records:

Pedigree	Petals		Bands	Colour	Margin	Wrinkling	Base		Remarks
	(*a*)	(*b*)					(α)	(β)	
144 30·16 826 23·3 C. F.	4	0	12	*p*	*n*	*sw*	*d*	*w*	Pollen brown

The first number (144) in the first column refers to the original Hampden great-grandmaternal plant, the second (30·16) to the Kidderminster I. grandmaternal capsule, the third (826) to the Parkstone maternal capsule, and the last number is the key number of the Crewe crop. C. F. signifies that it was from a free bud, or 826 capsule was in all probability cross-fertilised.

Again :

Pedigree	Petals (*a*)	Petals (*b*)	Bands	Colour	Margin	Wrinkling	Base (*α*)	Base (*β*)	Remarks
149 26·15 D. 47·1	6	3	13	*r*	*b*	*n*	*dl*	*w*	Conical topped capsule

This gives a grand-daughter capsule from Arncliffe of an original Hampden plant. It will be seen that with the crop-books before one it is possible to rapidly look up the pedigree of any plant, although the pollen ancestry must remain unknown.

By observing characters on the apical capsules only we at first considered that we should get rid of the difficulty of having to allow for homotyposis, i.e. the variability of character within the individual plant. This variability, as was pointed out in the first memoir, is very considerable. But as the work went on the question was again raised as to whether the apical flower is as true a measure of individuality as the totality of flowers on the plant. This point has been to some extent discussed in the first memoir on the Shirley*, and is by no means easy to answer. We found in that paper that offspring from seed of all the capsules of a plant had a greater average resemblance to the plant than offspring from seed taken from a single capsule, while the offspring were more like each other when their seed was taken from a single capsule. If the arrangement of buds on a plant were as definite as the digits of a man, we should have no hesitation in comparing the characters of the apical flowers of mother and daughter plants. But this is certainly not the case, and when we deal with characters which are by no means peculiar to the apical flower it becomes a very difficult point to determine how far the variation within the individual plant can or must be allowed for. Characters like "wrinkling," number of petals and number of stigmata have a very considerable variation within the individual. Frequently when the general tendency of the flowers of the plant is to a high number of stigmata, or to an abnormal number of petals, this tendency may fail to be represented on the first flower. Other characters like colour and margin may possibly vary less from flower to flower of the same plant; but are nevertheless sensibly influenced by environment. Lastly, a very considerable selection takes place of the capsules which give any germinating seed. The mother plants which provide germinating seed are far from a random sample of the general Shirley population.

* *Biometrika*, Vol. II. pp. 73, 74, and p. 82.

We have then the following points which must always be borne in mind and which render in our experience the problem of inheritance in plants extremely complex.

(*a*) The wide-reaching effect of a changed environment, whether this be a change of climate, soil, or season.

(*b*) The effect of selection in modifying the intensity of inheritance; the variability of the actual mother plants is for a number of characters far less than the variability in the general population.

(*c*) The difficulty as to the multiplicity of like organs in the individual and the choice as to which of those organs is to represent the individual.

We propose to consider these points in the following sections.

(3) *On the Influence of Environment on the Shirley Characters.*

In the first place let us consider the type and variability in number of stigmatic bands—a character which we can compare with the previous crops. Of course this comparison can only be partial, because in the earlier crops this character was recorded for every flower on the plant, while in 1903–4 only the first or apical flowers were dealt with. Further, for the earlier crops the mean and variability are those found for the homotyposis tables, while in the 1903–4 crops they are taken from the fraternal correlation tables.

TABLE I.

Change of Type and Variability with Environment.

Crop	Mean No. of Stigmata	S. D.	C. of V.
Parkstone, 1903 ...	14·03	1·562	11·13
Enfield I, 1900	13·78	1·518	11·02
Kidderminster II, 1903 ...	13·58	1·880	13·84
Crockham, 1900	13·39	2·055	15·35
Kidderminster I, 1900 ...	13·18	1·702	12·91
Bookham, 1900	13·14	2·330	17·73
Crewe, 1904	13·06	1·687	12·92
Enfield II, 1900	12·72	1·717	13·50
Hampden, 1899	12·61	1·885	14·95
Chelsea, 1899	12·37	1·680	13·58
Arncliffe, 1903	12·12	2·112	17·43
Chelsea II, 1900	12·04	1·694	14·07
Highgate, 1900	12·03	1·890	15·71
Oxford, 1900	11·39	2·173	19·08
Streatham, 1903	11·39	1·507	13·20

Considering that the probable error in these cases is a small quantity in the first or second place of decimals, we reach a very wide range of difference in type and variability occurring in different localities and seasons. We see at once that plants growing under diversities so great as are evidenced above are not directly comparable.

Without laying too much stress on this table we may remark that the mean variation of the first seven entries is 1·82, and of the last eight 1·83. In other

words, it does not appear that the circumstances which lower the mean of the stigmata are the same as those which lower variability. If we classify in order of variability we find the seven most variable stations give a mean of 12·6 bands, and the eight least variable a mean of 12·8 bands, two values, the difference of which is only doubtfully significant. In fact, it will be found that very fertile crops like Kidderminster I., Enfield I., and Parkstone rank in the same class of lesser variability with starveling crops like those of Streatham and Chelsea; while the high variability group embraces low types like those of Highgate, Arncliffe, and Oxford. The high situation, cold and late season of Arncliffe reduced the number of stigmata, but left the plants highly variable, the potting at Oxford produced the same result, but the causes which produced the starveling crop at Streatham were as influential in reducing its variability as in lowering its type. The general inference to be drawn is that both in variability and type the influence of environment, local or seasonal, on plant life is extremely large, and that careful investigations of the correlation of locality, soil, and seasonal meteorological conditions with plant type and variability would give valuable results. Meanwhile we can only adopt methods of studying resemblance, which as far as possible eliminate this influence.

TABLE II.

Percentages of Certain Characters in Different Crops of same Season.

Character	Crop	Percentage
With more than 13 Stigmata	Parkstone	63·0
" " "	Kidderminster II.	53·3
" " "	Arncliffe	24·8
" " "	Streatham	8·8
With more than four Petals or with some petaloid stamens ...	Parkstone	17·6
" " " " ...	Kidderminster II.	16·8
" " " " ...	Arncliffe	16·2
" " " " ...	Streatham	13·8
Percentages of Flowers with some red ...	Kidderminster II.	35·0
" " " " ...	Arncliffe	33·8
" " " " ...	Parkstone	25·5
" " " " ...	Streatham	25·0
Without margins	Kidderminster II.	72·6
" "	Parkstone	71·3
" "	Arncliffe	63·9
" "	Streatham	61·0
With slight or no basal patches	Streatham	52·4
" " "	Kidderminster II.	36·4
" " "	Parkstone	30·7
" " "	Arncliffe	18·4
With slight or no wrinkling	Parkstone	92·9
" " "	Kidderminster II.	78·7
" " "	Streatham	78·6
" " "	Arncliffe	65·6

Another suggestive manner of examining the influence of environment is to consider the percentage of certain characters observed in the different crops of one and the same season. We first place all the characters together without endeavouring to analyse how far the results may be due to differences of standard in different observers.

Now examining this table we see that although the seed for these four crops was all taken from a single local crop by sampling fertile individuals at random, yet the environment influenced widely the results. This is clear, because in counting the number of stigmata or noting if margins existed there was no room for sensible influence of personal equation. Luxuriant crops like those at Parkstone and Kidderminster had more stigmata, petals, and petaloid stamens than a poor crop like that at Arncliffe, or a starveling one like that at Streatham. Further we see that nearly the same order is preserved for the percentages without margins, or we suspect that poverty in a crop increases the number of individuals in whom the pigment does not reach the edge of the petal. With the other characters the results are not so congruous. Arncliffe and Streatham, which agree in having petals well margined, are different in basal patch frequency, Streatham showing more flowers without basal patch than any other crop and Arncliffe fewer. In wrinkling the results are not very definite, but indicate that the more luxuriant crops have less wrinkling; while reds appear to be favoured by the bleaker north. The chief point to be noticed is that there is immense diversity from one crop to a second, and that this diversity is hardly greater in characters like colour and base, where personal equation may be influential, than it is in stigmata and number of petals, where it must be of small or no account. All the above percentages have been deduced from the results weighted with the number of sister plants, so as to bring out the influence of change of environment on the results we have to work with in measuring sibling resemblance. These great differences between local crops of the same season forcibly suggest that individual plant-environment within the local crop, depending on subtle and inappreciable differences in soil, aspect, and amount of moisture, will in the case of material thus susceptible produce very sensible differentiations. We believe that a good deal of the irregularity, and possibly the lessened intensity of the inheritance coefficients for plant life, is due to this marked susceptibility of plants to small differences of environment. But until crops can be grown year by year in the same place, and under nearly similar conditions, it will be impossible to measure these conditions, or accurately determine their influence on the intensity of inheritance.

(4) *The Influence of Selection in Modifying the Intensity of Inheritance.*

If ρ be the actually observed correlation, and r the correlation if all members of a race produce offspring, then it is well known that any selection of parents, natural or artificial, tends to reduce ρ below the value of r. If the selection be stringent, then ρ may be so much lower than r that it may appear to have no relation to it at all, and fallacious conclusions be drawn as to the variability of inheritance for different species or for different characters in the same species,

the real fact of the case being that these species or characters are being subjected to very different intensities of selection.

If the distribution of the character be approximately normal, and if μ be the ratio of the variability of selected parents to the variability of the general population for a given character, then it is known that the true value of inheritance r is closely given by:

$$r = \frac{\rho}{\sqrt{1-(1-\mu^2)(1-\rho^2)}}.$$

Thus with anything like stringency of selection ρ may be quite small, and the value of r completely screened. We may illustrate this on various characters.

(*a*) *Stigmata.* In the case of a crop under very favourable conditions, or where every plant is separately grown as at Oxford, we have a high variability, amounting to 2 or over. But at Parkstone the actual variability was only 1·56, and on a random selection of 129 seed capsules from this crop only 92 sets germinated at Crewe, and of these sets the maternal plants had a stigmatic band variability of 1·4590. Thus $\mu = \cdot 7295$ for the stringency of selection of the maternal plants at Crewe, as compared with Shirley Poppies at large.

(*b*) *Extent of Margin.* If we take the 971 plants at Parkstone and the 408 offspring at Crewe from certain of them we note the following distributions (see p. 417):

	b	*bs*	*s*	*sn*	*n*	Totals
Parkstone General Population ...	52	73	84	71	691	971
Parkstone Mothers of Crewe Daughters ...	1	58	64	10	275	408

Or, dividing into three groups, we have the percentages:

	Broad Margins	Slight Margins	Insignificant
General Population ...	12·9	8·7	78·4
Mother Plants ...	14·5	15·7	69·8

We see that there has been a considerable alteration in the distribution of margins, between the actual mothers and the general population from which they were selected at random. Let us treat the distribution of margin as approximately normal to get some idea of the stringency of selection. In this case if d be the range of margins classed by this *one* observer as slight we find, if σ be the standard deviation of the general Parkstone population and σ' that of the selected mothers:

$$d = \cdot 3440\sigma \text{ and } d = \cdot 5497\sigma'.$$

Hence

$$\mu = \sigma'/\sigma = \cdot 62589.$$

Thus a fairly stringent selection of margins took place between the Parkstone general population and those members of it who actually provided germinating seed, and so daughter-plants at Crewe. But this is not the whole selection. Parkstone has for stigmatic bands a low variability, as compared with what we know of Shirleys in general. It is extremely unlikely that it has not also a reduced variability in extent of margin. How much, however, it is impossible to say, for we cannot compare Mr Blanchard's estimate of what is classed as "slight" in margin with estimates on other crops by other observers. We have no knowledge of the relative personal equations. Thus the above stringency of selection can only be considered as a minimum. μ in reality is probably considerably lower.

(*c*) *Number of Petals.* If we consider number of petals we find:

	Parkstone General Population	Parkstone Mothers of Crewe Crop	Kidderminster II. General Population
4 Petals	758	333	798
4 Petals and some petaloid stamens	89	21	47
More than 4 Petals with and without petaloid stamens	125	41	122
	972	395	967

Or, reduced to percentages:

4 Petals	78·0	84·3	82·5
4 Petals and some petaloid stamens	9·2	5·3	4·9
More than 4 Petals with and without petaloid stamens	12·8	10·4	12·6

Taking d as the range from normal four petals to the end of the range of normal four petals accompanied by petaloid stamens we have, if σ_P, σ', and σ_K be the respective standard deviations:

$$d = \cdot 3614\sigma_P = \cdot 2613\sigma' = \cdot 2090\sigma_K.$$

We see that in this case the Parkstone maternal plants are an "inverse" selection of the general population, being more variable than the general Parkstone crop, but they are considerably less variable than a typical crop like the Kidderminster. Thus on the general population of Shirleys there is a stringency of selection in the maternal plants of the Crewe crop at least equal to $\mu = \cdot 8$ and possibly more intense. At Streatham even the percentages being 86·1, 2·3, and 11·6 respectively, we have $d = \cdot 1103\sigma_S$, and the stringency of selection of the Crewe maternal plants would give μ as low as ·53. But there are all degrees of petaloid stamens, and all observers may not have noted equally every slight deviation from normality in the stamens of their flowers.

(*d*) *Colour of Middle Third of Petal.* We have the following results:

	Parkstone General Population	Parkstone Mothers of Crewe Crop
Some red in Petal*	240	153
All Pink ...	481	161
Some White ...	252	94
	973	408

Or, in percentages:

Some red in Petal	24·7	37·5
All Pink ...	49·4	39·5
Some White ...	25·9	23·0

Taking the all pink as our reference group we have

$$d = 1{\cdot}0562\sigma' = 1{\cdot}3315\sigma,$$

or

$$\mu = \sigma'/\sigma = 1{\cdot}2606.$$

In other words, there has been such a selection of red mothers that the mothers are *more variable* than the general population at Parkstone. The mean poppy, which in the general population is about half-way † up the pink range, has in the maternal plants crept nearly up to 1/8 from the beginning of the red group‡.

If we turn to the Crewe offspring we find the following values:

Mothers		Daughter Plants	
153	37·5 p.c.	139	34·1 p.c.
161	39·5 p.c.	199	48·8 p.c.
94	23·0 p.c.	70	17·1 p.c.
408		408	

Thus we see regression towards the pinks. In terms of d the pink range we have for σ'' the standard deviation of offspring

$$\sigma'' \times 1{\cdot}3586 = d,$$

or $\sigma/\sigma'' = 1{\cdot}02$ nearly. Thus the variability of the general population at Crewe is sensibly the same as that of the general population at Parkstone, but the mothers who provided germinating seed at Crewe were redder than the average Parkstone poppy, and the resulting offspring, though still regressing slightly towards the red, are yet about $\frac{1}{3}$ off the boundary of the red §, instead of half-way across the pink range as in the Parkstone crop.

The exact interpretation of this inverted selection is difficult, and has not yet been theoretically worked out. The formula on p. 404 involves approximate normality of the distribution, and becomes meaningless when the selection curve does not theoretically fall *wholly* inside that of the population from which selection takes place. The distributions of colour percentages at Kidderminster and Arncliffe are:

	Some Red	All Pink	Some White
Kidderminster ...	34·5	40·0	25·5
Arncliffe ...	33·8	40·4	25·8

* The two groups include under red 14 and 2 flowers respectively, which were noted to have a certain amount of white with the red or red-pink.

† ·6851 to 1·3315. ‡ ·3186 to 1·0562. § ·4106 to 1·3586.

and thus much alike, and very close to the Parkstone mothers of the Crewe crop. It seems therefore impossible to assert that the maternal plants at Crewe, although certainly an "inverse" selection from the Parkstone crop, are really a selection from the general population of Shirleys. We can only consider them as an average sample.

(*e*) *The Basal Patch.* With the sole exception of wrinkling the basal patch is the most unsatisfactory of the characters observed. Not only do the flowers showing no basal patch vary very much in percentage with the environment, but there is undoubtedly a considerable amount of divergence in the estimations made by different observers of the categories. We have the following results:

	Parkstone General Population	Parkstone Mothers of Crewe Crop	Crewe Crop
n, *ns*, *s* = Slight or none	295	134	53
sd, *d* = Definite ...	342	128	214
dl, *l* = Large ...	332	137	132
	969	399	399

Or, in percentages:

	Parkstone General Population	Parkstone Mothers of Crewe Crop	Crewe Crop
Slight or none	30·4	33·6	13·3
Definite ...	35·3	32·1	53·6
Large ...	34·3	34·3	33·1

Thus we see that while the maternal plants of the Crewe crop were a quite reasonably random sample of the Parkstone population, the distribution of Crewe offspring is totally different. How far is this due to environment, and how far to personal equation? The corresponding percentages at Kidderminster, Arncliffe, and Streatham are:

	Kidderminster General Population	Arncliffe General Population	Streatham General Population
Slight or none	36·5	18·4	52·4
Definite ...	50·4	47·7	30·9
Large ...	13·1	33·9	16·7

If these greater differences are not due to personal equation, which series can be taken as a typical crop of Shirleys? We confess that there does not seem to us any means of separating out the environmental influence from the personal equation in this case. All we can say is that there is no evidence that the maternal plants at Crewe were a selection from the general population of Shirleys, and if they were, we have no means of measuring the stringency of that selection.

(*f*) *Wrinkling of Petals.* This, which is a marked feature of the beauty of the Shirley, is the most unsuccessful of all the characters recorded. The instruction to observe it between 12 and 24 hours after opening of the bud could not always

be carried out, and there is little doubt that it is not only a feature much modified by season and environment but even by the interval between opening of the bud and the time of record.

The following table gives our results:

	Parkstone General Population	Parkstone Mothers of Crewe Crop	Crewe General Population
Frilled and Wrinkled ...	70	12	137
Slight Wrinkling ...	346	112	234
No Wrinkling ...	553	275	28
	969	399	399
Or, as percentages:			
Frilled and Wrinkled ...	7·2	3·0	34·3
Slight Wrinkling ...	35·7	28·1	58·7
No Wrinkling ...	57·1	68·9	7·0

Now it will be clear that while the maternal plants of the Crewe crop are a selected group of the Parkstone crop, yet there is no relation between the percentages in these mothers and in the offspring. The 70 per cent. of not wrinkled maternal plants actually produced four times as many wrinkled and eight times as many slightly wrinkled as not wrinkled plants, and the 3 per cent. wrinkled produced twice as many not wrinkled and slightly wrinkled as wrinkled plants!* In fact while the Parkstone crop had more non-wrinkled poppies than any other and the maternal selection 12 per cent. more than the general Parkstone population, the Crewe offspring had only 7·0 per cent. of non-wrinkled poppies, less than half that shown by the most wrinkled crop, Arncliffe, which showed about 15 per cent. Whatever changes took place in environment and season between Parkstone and Crewe they were sufficient to totally upset wrinkling as an inherited character! Some possibly, but far from all, of this difference we believe may be due to personal equation. The main feature to be noticed is that whereas the southern crops, Parkstone and Streatham, have more than 45 per cent. of unwrinkled flowers, the northern crops have less than 17 per cent. (Arncliffe 14·9, Kidderminster 16·4), sinking to 7 per cent. at Crewe.

If we confine our attention to the Parkstone crop and take d for the range of slightly wrinkled we have:

$$d = 1{\cdot}3860 \times \sigma' = 1{\cdot}2812\,\sigma,$$

Or the stringency of maternal selection

$$= \mu = \sigma'/\sigma = {\cdot}9244.$$

If we take the Oxford crop of 1900 as one in which the wrinkling was very carefully observed, we find $d = 1{\cdot}1946\,\sigma$, or we have $\mu = \sigma'/\sigma = {\cdot}8620$; so that it is quite possible that the stringency may be greater than that indicated by the Parkstone crop.

* See Appendix I., Table F.

(5) *The Influence of Variability of like Organs in the Plant on Coefficients of Heredity. Homotyposis.*

We have already seen that if all the flowers of a plant were identical, or if the first flower were as markedly differentiated from other flowers as one digit in a man is from a second (i.e. if there were a very marked correlation between order of flowering and character), the problem of inheritance in plants would be much simpler. But the variation of like organs in a plant is very considerable, and in many cases the degree of variation is not highly correlated with the position or order of production. Now if there were no such correlation, it has been shown elsewhere that the extent of the variation of the individual plant—its homotyposis—would be properly measured by the coefficient of correlation of like organs. Hence our selection for record of apical flowers will mean that we get closer to the true values of the inheritance coefficients than by using all or any flowers on the plant, provided 1st there be marked differentiation of the apical flower from other flowers (i.e. if the first flowers of two related plants are much more alike than the first flower of one and any other flower of the second), or again 2nd there be no variation of the character observed in the plant at all. For example, colour and margin in the broad classification of our records do not exhibit much variation in the plant. Number of stigmata, size of local patch and wrinkling, do exhibit much variation, and further this variation is by no means perfectly correlated with position or order of budding. In order therefore to allow for the neglected homotyposis, we ought to have obtained measurements of (i) the correlation between the order of flowering and the characters observed, and (ii) the homotyposis of these characters. The first would have enabled us to measure the differentiation of the first or apical flower from other flowers, and by the aid of the second result we could have made allowance for the neglected homotyposis.

But to record some half-dozen characters on all the flowers of each plant would have multiplied the work of recording tenfold, and the task seemed at present beyond the volunteer labour, which was all that was available. Accordingly we must simply state that our inheritance coefficients are too low because we have at present no adequate measure of the differentiation of the apical flowers, nor of the homotyposis of most of the characters dealt with. But it is satisfactory to note that those characters in which mere inspection has shown least differentiation with position and most variation in the individual, such as number of stigmata, size of basal patch and wrinkling, are precisely those in which the coefficients of inheritance come out lowest. We hope that investigations into character differentiations with position and into homotyposis will be started this season.

Meanwhile the results in sections (3) to (5) will, we hope, suffice to indicate to the reader that in our opinion the study of inheritance in plants is not a simple, straightforward matter. We have not in Shirleys at any rate found the definite and exclusive categories which the Mendelians appear to encounter at every turn. We have found our characters widely modified by environment and season; we have found selection acting on the seed sown, and both type and variability

susceptible of much alteration. Lastly, although we have classified each character into more sub-groups than we have met with in Mendelian writings, we find many cases in which the category is hard to determine, and we are convinced that personal equation must be given its due weight.

(6) *Parental Correlation.*

As we stated at the beginning of this paper, we intended to base the 1903 experiments on examinations of apical flowers only. We were not therefore able to compare in the bulk of characters the crops of 1903 with the maternal plants. Ample material should, however, have been provided for parental correlation in the three crops of 1904. Unfortunately two of our observers were unable to complete their tasks. We are therefore thrown back on a comparison of the Crewe daughter plants with the Parkstone maternal plants. The fact that the Parkstone and Crewe crops were grown in very different environments is against very good results, but the number of bagged capsules at Parkstone and their comparatively good state led to our selecting that crop as the chief source for the seed supply of 1903. The tables are formed for the apical flowers of the maternal plants at Parkstone and the daughter plants at Crewe, and this was done with six characters. The work of tabling and reduction is due to Dr A. Lee. The coefficients of correlation, except in the case of the stigmata where the method of moments was used, were found by the fourfold table method. We should, if the investigations were now to be started *de novo*, probably use the method of mean square contingency, but it was considered desirable to preserve as far as possible uniformity of treatment throughout the whole round of poppy investigations, and the fourfold table method had been adopted in the first memoir. We refer briefly to the characters dealt with.

Stigmata. There was probably a stringency of selection here of the maternal plants of about ·7295. It is an open question to what extent the homotyposis of about ·5 ought also to be allowed for, at any rate the value found for parental inheritance is a minimum value.

Extent of Margin. This was determined by a fourfold table, the divisions being between 'no' and 'some' margins. The factor of selection of the maternal plants is here ·6259. The homotypic influence is, as in the case of colour, probably very small, the flowers on a plant being usually all whole coloured or all margined.

Number of Petals. This was determined by a fourfold table, division into normal four petals and non-normal flowers. There was inverse selection of maternal plants compared to the general Parkstone population, i.e. they were more variable than that population; but if we compare the maternal group with the Streatham crop as a Shirley population of maximum petal variability there would have been a stringency of selection of ·5278. It is not clear, however, how great a selection we ought to allow for, because the observers may not have been equally definite as to what they reckoned as an incipient petaloid

stamen. Further, we have at present no measure of the homotyposis of plants in relation to supernumerary petals and petaloid stamens.

Colour of Middle Third of Petal. This was worked by a fourfold table, the division being into poppies showing some and those showing no red. We cannot do more than consider the maternal plants as an average random selection of Shirleys (see p. 406). In the matter of body colour, within the classifications red, pink, pink-white, etc., there appears to be small variation within the plant, or we may look upon this broad colour classification as individual to the plant and not to the bud.

The Basal Patch. This was dealt with by a fourfold table, the division being, none to slight, and slight-definite to large. No evidence is available as to whether there was or was not a selection of maternal plants. The extent of basal patch varies considerably on flowers of the same plant, and is probably a good deal influenced by personal equation and homotyposis.

Wrinkling. The division here was by a fourfold table. Mr Blanchard's estimates of wrinkling were so low as compared with Mr Ward's that the division of the maternal plants was made into not wrinkled and some wrinkling, while the Crewe population was divided into wrinkled and slight or no wrinkling. There may be a selection stringency of ·8 to ·9, and probably normal homotyposis, as wrinkling varies considerably on the same plant. But we have no definite knowledge on these points, and we consider wrinkling, although very characteristic of the Shirley, as very unsuitable for record.

We sum up our results in the following table:

TABLE III.

Parental Correlation, Parkstone Maternal and Crewe Daughter Plants.

Character	Raw Correlation	Selection Stringency	Corrected Correlation
Stigmata...	·1717	·7295	·2324
Extent of Margin ...	·3175	·6259	·4718
Number of Petals ...	·2740	·5278 (?)	·4750 (?)
Colour of Middle Third	·4008	0	·4008
Basal Patch	·1305	·6900 (??)	·1864 (?)
Wrinkling	·1595	·8620 (??)	·1842 (?)
Mean	·24	·69	·33

We have not the least hesitation in admitting the full inconclusiveness of these results. All we consider that they show is, that if we could accurately allow for selection of parentage, for the variability of character in a plant with its multiplicity of organs, and for the extreme response to seasonal and environmental changes, and, further, eliminate the effects of personal equation, there is no reason to suppose that the parental resemblances for different characters would

not group themselves between ·4 and ·5, the values found for characters in other types of life. We do not believe that experiments on plant heredity, which employ no measurement and allowance for these modifying factors, can lead to conclusive results. In our opinion these factors are all potent, and modify widely the limits of any category which may be selected for classification of simple characters in plants.

(7) *On Collateral Heredity.*

The following table gives the results obtained for the 1903 crops with a comparison of such characters as are comparable in the Oxford crop of 1900. In all cases the apical capsules in the sibling plants are compared. In the Appendix are given the actual tables from which the results are deduced. It is desirable to point out the methods adopted in the individual cases:

(*a*) *Resemblance in number of Stigmatic Bands.* Found by usual correlation table.

(*b*) *Number of Petals.* Fourfold table, the grouping being into the normal four petals and the abnormal more than four petals, including in the latter class semi-double, double flowers, and flowers with one or more petaloid stamens.

(*c*) *Colour of Middle Third.* Fourfold table, the grouping being into flowers showing some and flowers showing no red.

(*d*) *Breadth of Margin.* Fourfold table, the grouping being into flowers showing some and flowers showing no margin.

(*e*) *Colour of Flower at Margin of Petal.* Fourfold table, the grouping being into flowers showing some and flowers showing no white there.

(*f*) *Extent of Basal Patch.* Fourfold table, the divisions being into none or slight, and large or definite patches.

(*g*) *Wrinkling.* Fourfold table, the divisions being into marked wrinkling ('frilled' or 'wrinkled') and insignificant wrinkling (slight or no wrinkling)*.

Several points are brought out by this table. In the first place we see that the starveling Streatham crop shows a great reduction in the fraternal correlations. Now if we endeavour to interpret this result we should reasonably suppose that much selection took place at Streatham; the mean of the stigmatic bands and their variability has been much lowered. There is further no reason for supposing a correlated selection in brother plants. Now if we suppose two organs A and B with standard deviations σ_1 and σ_2 and correlation r_{12}, to be selected in a non-correlated manner so that their standard deviations become s_1 and s_2, it is easy to show, on the hypothesis of nearly normal distribution, that the correlation after selection will be

$$\rho_{12} = \frac{s_1 s_2 r_{12}}{\sigma_1 \sigma_2 \sqrt{1 - r_{12}^2 \left(1 - \frac{s_1^2}{\sigma_1^2}\right)} \sqrt{1 - r_{12}^2 \left(1 - \frac{s_2^2}{\sigma_2^2}\right)}} \quad \ldots\ldots\ldots \text{(i)},$$

and this independently of changes in the means.

* Shape of the capsules, for which only data at Crewe were available, gave an uncorrected correlation between sister plants of ·15. See Table VV. Appendix.

TABLE IV.

Fraternal Correlation, Record of Apical Flowers.

Character	Crop					Mean
	Oxford, 1900	Kidderminster II, 1903	Arncliffe, 1903	Parkstone, 1903	Streatham, 1903	
Stigmata	·26	·33	·38	·14	·16	·25
Colour of Middle Third...	·34	·43	·38	·41	·31	·38
Breadth of Margin ...	·19	·29	·21	·15	·14	·19
Number of Petals ...	·23	·25	·16	·28	·17	·22
Extent of Basal Patch ...	·22	·24	·16	·18	·06	·17
Colour at Margin ...	—	—	—	·28	—	·28
Wrinkling	·21	·16	·12	·11	·06	·13
Petal Length	·28	—	—	—	—	·28
Mean	·24	·28	·23	·22	·15	·24

Accordingly, applying this to the case of two brothers, with a standard deviation in their generation of s, and an unselected standard deviation σ, we conclude that if a generation of any population be subjected to a selection of stringency $s/\sigma = \mu$, the correlation of brothers r will be reduced to a correlation ρ where

$$\rho = \frac{r\mu^2}{1 - r^2 + r^2\mu^2} \quad \text{..... (ii).}$$

Now if $r = \frac{1}{4}$ be the true correlation of half siblings, what must be the intensity of selection μ required to reduce this to $\rho = \cdot 15$?

Substituting, we at once find $\mu = \cdot 7645$.

Now can we suppose that a selection of this stringency has been at work at Streatham? It would involve a reduction of about 25 per cent. in the variability below the normal variability of the Shirley Poppy. Now the only data we can judge from are the first three crops which give a mean fraternal correlation of ·25; and the only character for which we know their absolute variability is their number of stigmatic bands. These give a mean variability of 2·055, which is practically equal to that in the general Shirley population. Accordingly the selected variability of the Streatham poppies ought to be $2{\cdot}055 \times {\cdot}7645 = 1{\cdot}571$. The actually observed value is 1·507. In like manner the Parkstone crop which shows the next greatest reduction in the fraternal correlation has for its reduced stigmatic band correlation ·14, which should mark a reduction in stigmatic band variability from about 2 to 1·57, the actual value observed being 1·56. Without laying too much stress on isolated accordance in a very irregular table of this kind we may reasonably conclude that the great apparent reduction in the Streatham correlations, and to a less extent the reduction in the Parkstone, can

most probably be accounted for on the basis of selection, such selection really reducing the correlations in a manner qualitatively and, as far as can be roughly determined, quantitatively in the manner observed.

In the next place, we have here made no allowance for possible selection of maternal plants and none for homotyposis. Hence the value of our fraternal correlation is again a minimum value. Even including the Streatham results the average correlation for all characters is ·23 and without these it is ·26. The chief irregularities are in the values of the correlations for wrinkling, extent of basal patch and margin, where we should expect minute differences of soil or situation, time of observation or personal equation to be more influential*. Now we might at first argue that ·25 would be a satisfactory value to reach for fraternal heredity, as being about *half* the value noted for fraternal correlation of whole brothers, but there are three points to be considered here. (i.) The neglect of any correction for selection, whether of maternal plants or of surviving offspring, has certainly given too low a value to these correlations. (ii.) There is a further correction to be made for the homotyposis which remains after allowing for differentiation of the first flower. This is probably not substantial in some of the characters, but it is considerable in others. It would again tend to raise the observed values. (iii.) In considering the result we have assumed that daughter plants grown from the same capsules are *half* sisters, or that the pollen parents are different for each pair of cases. We have already indicated (p. 398) that we consider that there is a fairly wide cross-fertilisation, i.e. a considerable variety of pollen parents to plants grown from the same capsule, but it is hardly likely to be complete for each pair of sister plants, and will probably vary considerably from capsule to capsule. Thus half-sibling correlation is rather a minimum limit to what we might anticipate, and not what is actually likely to occur.

Hence the apparent reasonableness of a value about ·25 is seen to be rather fortuitous.

Lastly, granted that the sister plants are in the bulk half-sisters, and that the average value of the correlation of whole sisters is about ·5, it appears open to considerable questioning whether we ought really to expect that the half-sibling resemblance will be just *half* the whole-sibling resemblance. This may appear to some to flow at once from *à priori* considerations, but we do not feel so clear on this point now as a few years ago; for the determinations of whole- and half-sibling relationships which have been made so fa" do not bear it out†.

* Experiments at University College on personal equation show a continuous change during a long series of observations. Judgment would thus depend on the time of flowering, and brother plants not being registered usually on the same occasion, it follows that personal equation would be likely to introduce a small spurious correlation between non-brothers, rather than to emphasise the degree of resemblance between brothers.

† As in the case of all vital problems, however, the complexity is very great. Horse and cattle breeders usually inbreed to some extent, and it very often happens that two half-siblings have more than four common great-grandparents, and so should be more alike than normal half-siblings with only four in common.

The following table gives a summary of the results hitherto reached: we have included the Shirley as if the sister plants were true half-siblings:

Half-Sibling Resemblance.

Thoroughbred Horse, through dam (3 series)	... ...	·36
Shorthorn Cattle, through sire and dam (6 series)	...	·32
Basset Hound, through dam (1 series)	... ...	·22
Shirley Poppy, Co-ovarial Plants.	Colour (5 series) ...	·38
" " "	All characters (32 series)	·24

In our opinion the horse and the shorthorn are the most reliable of these series, and in both cases the values are sensibly larger than half the whole sibling values. The Shirley when we deal with all the characters is fairly close to ·25 and the Basset Hound value. But the colour resemblance in the Shirley, which we consider least affected by homotyposis and selection, is in closer agreement to what we find for horses and cattle. We feel therefore that differential local environment, maternal and offspring selection, varying personal equation in estimating plant characters, and the complete absence at present of any quantitative allowance for positional differentiation and homotyposis, prevent us from saying more than that the results so far obtained for fraternal correlation in Shirleys are not inconsistent with those found for animals.

Much enlightenment would be reached could we deal with definite pollen parents in very large numbers. Further what is badly needed for our general guidance at the present time is a measure of the relative intensity of whole and half-sibling resemblance in some quantitatively determinable character. Possibly the record of the cephalic indices of whole and half brothers in public primary schools would be a comparatively easy method of reaching a definite result.

(8) *General Conclusions.*

(*a*) We consider that in colour distribution and tint as well as in other characters there is continuity in the variation of the Shirley Poppy.

(*b*) The Shirley Poppy under ordinary conditions of environment and season is cross-fertilised and probably there is a very considerable variety of pollen parent to sister plants grown from the seed of one capsule.

(*c*) The very great influence of environment upon nearly all the characters dealt with, the variation of character within the same plants, and the unknown allowance to be made for homotyposis, even when we confine our attention to apical flowers, render the problem of plant heredity of very great complexity.

(*d*) As suitable measurable characters in a plant are extremely few, one is compelled to select qualitative characters; and their description, especially when up to the present no scale of patterns has been available, is difficult, and liable to considerable modification by the influence of personal equation.

The experience of the present series of observations by no means accords with that of those observers who find it easy to formulate simple alternative categories

for plant characters, and group without apparent hesitation individuals into one or other category.

(*e*) The *minimum limit* to parental correlation has the average value ·33. In the best class of characters, however, it lies between ·4 and ·48, and we believe if the disturbing factors could be fully allowed for it would be found to lie between ·45 and ·50, its value as determined for other types of life.

(*f*) The *minimum limit* to fraternal correlation has the average value ·24. This allows for no selection of maternal plants. The real value is undoubtedly higher, and is quite consistent with half-sibling resemblance as found for horses and cattle, which it closely approaches in the case of colour, where the influence of selection and of variation in the plant appears to be least.

(*g*) Although very considerable differences exist in the values found for the correlations at different centres and for different characters, we do not think these differences are sufficient to justify any assertion that the intensity of heredity varies with character or with environment. They point rather to the great difficulty of dealing with plant life. What is urgently needed is a biometric farm where secular experiments could be conducted under nearly uniform conditions; there a continuous record of all flowers of each plant could be made by one and the same observer using a standard qualitative scale of groups.

APPENDIX. CORRELATION TABLES *.

I. RESEMBLANCE OF DAUGHTER AND MOTHER PLANTS. DAUGHTER PLANT GROWN FROM SEED TAKEN FROM APICAL CAPSULE OF MOTHER PLANT.

A. *Parkstone and Crewe. Apical Flowers.*

Parkstone Mother Plant. Number of Stigmata.

Crewe Daughter Plant. Number of Stigmata.	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*	*19*	Totals
9	—	—	—	1	1	—	—	—	—	—	2
10	1	4	4	6	3	4	—	—	—	—	22
11	—	6	6	5	12	7	1	—	—	—	37
12	—	4	13	20	26	13	2	—	—	—	78
13	1	10	16	23	29	14	2	3	—	—	98
14	—	3	15	18	22	13	2	3	2	—	78
15	—	1	6	13	15	11	3	2	1	—	52
16	—	4	2	1	10	6	2	3	—	—	28
17	—	1	—	—	2	1	1	—	—	—	5
18	—	1	—	1	—	—	—	—	—	—	2
19	—	—	—	—	2	—	—	—	—	—	2
Totals	2	34	62	88	122	69	13	11	3	0	404

* There is not absolute uniformity in the classification categories of all these tables. Thus one observer may have used "wrinkled" entirely and never the extreme "frilled," while a second may have placed the few "veined whites" in a category denoting white with very little pink.

B. *Parkstone and Crewe. Apical Flowers.*

Parkstone Mother Plant. Number of Petals and Petaloid Stamens.

Crewe Daughter Plant. Number of Petals and Petaloid Stamens.	4 Petals	4 Petals + Petaloid Stamens	5 Petals or more without and with Petaloid Stamens	Totals
4 Petals	261	12	26	299
4 Petals + Petaloid Stamens ...	54	7	14	75
5 or more Petals without and with Petaloid Stamens	18	2	1	21
Totals	333	21	41	395

C. *Parkstone and Crewe. Apical Flowers.*

Parkstone Mother Plant. Colour of Middle Third of Petal.

Crewe Daughter Plant. Colour of Middle Third of Petal.	*r*	*rp*	*rw*	*rpw*	*p*	*pw*	*vw*	*w*	Totals
r	37	23	—	—	29	3	1	1	94
rp	4	11	—	—	18	9	1	1	44
rw	—	—	—	—	—	—	—	—	—
rpw	—	1	—	—	—	—	—	—	1
p	31	32	—	—	94	32	5	5	199
pw	2	6	1	1	17	20	6	6	59
vw	—	—	—	—	—	—	1	—	1
w	1	3	—	—	3	2	1	—	10
Totals	75	76	1	1	161	66	15	13	408

D. *Parkstone and Crewe. Apical Flowers.*

Parkstone Mother Plant. Breadth of Margin.

Crewe Daughter Plant. Breadth of Margin.	*b*	*bs*	*s*	*sn*	*n*	Totals
b	1	18	12	3	20	54
bs	—	4	7	—	18	29
s	—	2	3	—	8	13
sn	—	11	6	3	42	62
n	—	23	36	4	187	250
Totals	1	58	64	10	275	408

E. *Parkstone and Crewe. Apical Flowers.*

Parkstone Mother Plant. Extent of Basal Patch.

Crewe Daughter Plant. Extent of Basal Patch.	*n*	*ns*	*s*	*sd*	*d*	*dl*	*l*	Totals
n	2	2	1	2	—	—	2	9
ns	1	1	1	3	—	3	—	9
s	4	3	7	9	7	4	1	35
sd	1	—	1	3	1	4	1	11
d	28	3	20	32	37	60	23	203
dl	—	—	—	—	—	—	—	—
l	38	6	15	21	13	24	15	132
Totals	74	15	45	70	58	95	42	399

F. *Parkstone and Crewe. Apical Flowers.*

Parkstone Mother Plant. Wrinkling of Petal.

Crewe Daughter Plant. Wrinkling of Petal.	*f* or *w*	*sw*	*nw*	Totals
f or *w*	4	47	86	137
sw	5	60	169	234
nw	3	5	20	28
Totals	12	112	275	399

II. RESEMBLANCE OF SISTER PLANTS, *i.e.* PLANTS GROWN FROM SEED TAKEN FROM THE SAME MATERNAL CAPSULE.

AA. *Parkstone. Sister Plants' Apical Flowers.*

First Sister Plant. Number of Stigmata.

Second Sister Plant. Number of Stigmata.	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*	*19*	Totals
10	2	3	18	12	30	13	4	6	—	—	88
11	3	26	38	53	51	27	18	11	—	1	228
12	18	38	100	167	184	132	63	27	4	2	735
13	12	53	167	310	315	278	115	51	7	4	1312
14	30	51	184	315	404	332	168	80	15	5	1584
15	13	27	132	278	332	318	168	73	15	1	1357
16	4	18	63	115	168	168	84	47	13	2	682
17	6	11	27	51	80	73	47	20	7	2	324
18	—	—	4	7	15	15	13	7	—	—	61
19	—	1	2	4	5	1	2	2	—	—	17
Totals	88	228	735	1312	1584	1357	682	324	61	17	6388

BB. *Arncliffe. Sister Plants' Apical Flowers.*

First Sister Plant. Number of Stigmata.

Second Sister Plant. Number of Stigmata.	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*	*19*	Totals
3	—	—	—	—	—	—	—	1	2	3	2	1	—	—	—	—	—	9
4	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
5	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0
6	—	—	—	2	2	4	4	4	2	—	—	—	—	—	—	—	—	18
7	—	—	—	2	—	4	3	5	4	3	3	3	—	1	—	—	—	28
8	—	—	—	4	4	18	20	33	19	18	13	7	1	—	—	—	—	137
9	—	—	—	4	3	20	16	51	44	35	17	5	7	1	—	—	—	203
10	1	—	—	4	5	33	51	116	127	119	90	54	13	7	—	—	—	620
11	2	—	—	2	4	19	44	127	198	136	138	92	37	37	1	1	—	838
12	3	—	—	—	3	18	35	119	136	230	168	132	57	30	2	4	—	937
13	2	—	—	—	3	13	17	90	138	168	134	126	45	29	6	9	2	782
14	1	—	—	—	3	7	5	54	92	132	126	104	38	36	9	1	—	608
15	—	—	—	—	—	1	7	13	37	57	45	38	26	31	3	8	2	268
16	—	—	—	—	1	—	1	7	37	30	29	36	31	20	13	6	1	212
17	—	—	—	—	—	—	—	—	1	2	6	9	3	13	6	1	—	41
18	—	—	—	—	—	—	—	—	1	4	9	1	8	6	1	6	3	39
19	—	—	—	—	—	—	—	—	—	—	2	—	2	1	—	3	—	8
Totals	9	0	0	18	28	137	203	620	838	937	782	608	268	212	41	39	8	4748

CC. *Streatham. Sister Plants' Apical Flowers.*

First Sister Plant. Number of Stigmata.

Second Sister Plant. Number of Stigmata.	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	Totals
6	—	—	—	1	2	1	—	—	—	—	—	—	4
7	—	—	1	3	9	14	9	6	1	—	—	—	43
8	—	1	6	14	54	38	55	27	11	4	—	—	210
9	1	3	14	90	199	234	224	72	28	5	3	—	873
10	2	9	54	199	436	580	568	198	91	27	8	1	2173
11	1	14	38	234	580	754	777	328	158	41	17	2	2944
12	—	9	55	224	568	777	734	362	199	50	19	4	3001
13	—	6	27	72	198	328	362	184	143	38	23	1	1382
14	—	1	11	28	91	158	199	143	70	24	12	1	738
15	—	—	4	5	27	41	50	38	24	2	1	1	193
16	—	—	—	3	8	17	19	23	12	1	6	—	89
17	—	—	—	—	1	2	4	1	1	1	—	—	10
Totals	4	43	210	873	2173	2944	3001	1382	738	193	89	10	11660

DD. *Kidderminster. Sister Plants' Apical Flowers.*

First Sister Plant. Number of Stigmata.

Second Sister Plant. Number of Stigmata.	*6*	*7*	*8*	*9*	*10*	*11*	*12*	*13*	*14*	*15*	*16*	*17*	*18*	*19*	Totals
6	—	—	—	—	2	—	2	1	—	2	2	—	—	—	9
7	—	—	—	—	3	3	1	2	—	—	1	—	—	—	10
8	—	—	8	20	17	3	2	2	1	—	—	—	—	—	53
9	—	—	20	26	23	7	6	7	10	5	2	—	—	—	106
10	2	3	17	23	34	34	39	44	41	17	11	5	1	—	271
11	—	3	3	7	34	56	101	113	131	58	22	5	6	—	539
12	2	1	2	6	39	101	236	285	280	158	94	23	8	7	1242
13	1	2	2	7	44	113	285	362	440	206	129	64	10	6	1671
14	—	—	1	10	41	131	280	440	522	289	174	109	34	10	2041
15	2	—	—	5	17	58	158	206	289	210	119	57	20	8	1149
16	2	1	—	2	11	22	94	129	174	119	92	59	8	2	715
17	—	—	—	—	5	5	23	64	109	57	59	64	14	3	403
18	—	—	—	—	1	6	8	10	34	20	8	14	2	—	103
19	—	—	—	—	—	—	7	6	10	8	2	3	—	—	36
Totals	9	10	53	106	271	539	1242	1671	2041	1149	715	403	103	36	8348

EE. *Parkstone. Sister Plants' Apical Flowers.*

First Sister Plant. Number of Petals and Petaloid Stamens.

Second Sister Plant. Number of Petals and Petaloid Stamens.	4 petals	4 petals + petaloid stamens	5 petals	5 petals + petaloid stamens	More than 5 petals	More than 5 petals + petaloid stamens	Totals
4 petals	5684	457	102	70	172	206	6691
4 petals + petaloid stamens	457	88	20	17	16	47	645
5 petals	102	20	8	2	9	9	150
5 petals + petaloid stamens	70	17	2	4	7	9	109
More than 5 petals	172	16	9	7	8	9	221
More than 5 petals + petaloid stamens	206	47	9	9	9	24	304
Totals ...	6691	645	150	109	221	304	8120

FF. *Arncliffe. Sister Plants' Apical Flowers.*

First Sister Plant. Number of Petals and Petaloid Stamens.

Second Sister Plant. Number of Petals and Petaloid Stamens.	4 petals	4 petals + petaloid stamens	Five petals, more than 5 petals, 5 or more with petaloid stamens	Totals
4 petals	3358	450	213	4021
4 petals + petaloid stamens	450	98	42	590
5 petals, more than 5 petals, 5 or more petals with petaloid stamens	213	42	32	287
Totals ...	4021	590	287	4898

GG. *Streatham. Sister Plants' Apical Flowers.*

First Sister Plant. Number of Petals and Petaloid Stamens.

Second Sister Plant. Number of Petals and Petaloid Stamens.	4 petals	4 petals + petaloid stamens	5 or more petals	5 or more petals + petaloid stamens	Totals
4 petals	8794	206	373	703	10076
4 petals + petaloid stamens	206	6	17	35	264
5 or more petals ...	373	17	26	47	463
5 or more petals + petaloid stamens	703	35	47	102	887
Totals ...	10076	264	463	887	11690

HH. *Kidderminster. Sister Plants' Apical Flowers.*

First Sister Plant. Number of Petals and Petaloid Stamens.

Second Sister Plant. Number of Petals and Petaloid Stamens.	4 petals	4 petals + petaloid stamens	5 or more petals	5 or more petals + petaloid stamens	Totals
4 petals	5868	268	241	505	6882
4 petals + petaloid stamens	268	48	15	74	405
5 or more petals ...	241	15	8	27	291
5 or more petals + petaloid stamens	505	74	27	90	696
Totals ...	6882	405	291	696	8274

II. *Parkstone. Sister Plants' Apical Flowers.*

First Sister Plant. Colour of Middle Third of Petal.

Second Sister Plant. Colour of Middle Third of Petal.	*r*	*rp*	*rw*	*rpw*	*p*	*pw*	*ppw*	*vw*	*w*	Totals
r	320	147	8	11	407	36	12	34	16	991
rp	147	286	15	11	458	81	18	49	14	1079
rw	8	15	—	1	21	1	—	1	4	51
rpw	11	11	1	6	44	3	2	6	1	85
p	407	458	21	44	2262	452	140	319	109	4212
pw	36	81	1	3	452	221	50	111	38	994
ppw	12	18	—	2	140	50	52	30	17	321
vw	34	49	1	6	319	111	30	76	26	652
w	16	14	4	1	109	38	17	26	26	251
Totals	991	1079	51	85	4212	994	321	652	251	8636

JJ. *Arncliffe. Sister Plants' Apical Flowers.*

First Sister Plant. Colour of Middle Third of Petal.

Second Sister Plant. Colour of Middle Third of Petal.	*r*	*rp*	*p*	*pw*	*vw*	*w*	Totals
r	390	155	300	109	5	53	1012
rp	155	126	217	74	4	38	614
p	300	217	986	315	10	114	1942
pw	109	74	315	254	7	84	843
vw	5	4	10	7	—	4	30
w	53	38	114	84	4	76	369
Totals	1012	614	1942	843	30	369	4810

KK. *Streatham. Sister Plants' Apical Flowers.*

First Sister Plant. Colour of Middle Third of Petal.

Second Sister Plant. Colour of Middle Third of Petal.	*r*	*rp*	*p*	*pw*	*pww*	*w*	Totals
r	792	142	992	89	208	164	2387
rp	142	36	242	28	35	39	522
p	992	242	4144	167	516	409	6470
pw	89	28	167	22	42	38	386
pww	208	35	516	42	148	87	1036
w	164	39	409	38	87	116	853
Totals	2387	522	6470	386	1036	853	11654

LL. *Kidderminster. Sister Plants' Apical Flowers.*

First Sister Plant. Colour of Middle Third of Petal.

Second Sister Plant. Colour of Middle Third of Petal.	*r*	*rp*	*p*	*pw*	*pww*	*rw*	*w*	Totals
r	714	251	473	170	49	11	59	1727
rp	251	308	381	155	47	—	39	1181
p	473	381	1392	557	181	5	103	3092
pw	170	155	557	414	104	2	83	1485
pww	49	47	181	104	46	—	50	477
rw	11	—	5	2	—	—	—	18
w	59	39	103	83	50	—	50	384
Totals	1727	1181	3092	1485	477	18	384	8364

MM. *Parkstone. Sister Plants' Apical Flowers.*

First Sister Plant. Breadth of Margin.

Second Sister Plant. Breadth of Margin.	*b*	*bs*	*s*	*sn*	*n*	Totals
b	46	51	44	33	279	453
bs	51	44	69	55	431	650
s	44	69	78	57	456	704
sn	33	55	57	82	450	677
n	279	431	456	450	4542	6158
Totals	453	650	704	677	6158	8642

NN. *Arncliffe. Sister Plants' Apical Flowers.*

First Sister Plant. Breadth of Margin.

Second Sister Plant. Breadth of Margin.	*b*	*bs*	*s*	*sn*	*n*	Totals
b	244	56	99	39	536	974
bs	56	10	9	21	87	183
s	99	9	54	18	209	389
sn	39	21	18	24	161	263
n	536	87	209	161	2070	3063
Totals	974	183	389	263	3063	4872

OO. *Streatham. Sister Plants' Apical Flowers.*

First Sister Plant. Breadth of Margin.

Second Sister Plant. Breadth of Margin.	*b*	*bs*	*s*	*sn*	*n*	Totals
b	136	87	205	107	624	1159
bs	87	42	93	59	340	621
s	205	93	290	165	885	1638
sn	107	59	165	112	686	1129
n	624	340	885	686	4582	7117
Totals	1159	621	1638	1129	7117	11664

PP. *Kidderminster. Sister Plants' Apical Flowers.*

First Sister Plant. Breadth of Margin.

Second Sister Plant. Breadth of Margin.	*b*	*bs*	*s*	*sn*	*n*	Totals
b	66	—	84	66	357	573
bs	—	—	—	—	—	0
s	84	—	228	116	607	1035
sn	66	—	116	88	416	686
n	357	—	607	416	4690	6070
Totals	573	0	1035	686	6070	8364

QQ. *Parkstone. Sister Plants' Apical Flowers.*

First Sister Plant. Extent of Basal Patch.

Second Sister Plant. Extent of Basal Patch.	*n*	*ns*	*s*	*sd*	*d*	*dl*	*l*	Totals
n	474	83	108	184	343	266	307	1765
ns	83	20	39	30	66	48	44	330
s	108	39	52	60	126	86	72	543
sd	184	30	60	144	240	160	100	918
d	343	66	126	240	582	453	293	2103
dl	266	48	86	160	453	354	236	1603
l	307	44	72	100	293	236	274	1326
Totals	1765	330	543	918	2103	1603	1326	8588

RR. *Arncliffe. Sister Plants' Apical Flowers.*

First Sister Plant. Extent of Basal Patch.

Second Sister Plant. Extent of Basal Patch.	*n*	*ns*	*s*	*sd*	*d*	*dl*	*l*	Totals
n	156	13	28	36	237	140	37	647
ns	13	2	1	6	20	12	2	56
s	28	1	8	17	73	52	15	194
sd	36	6	17	34	125	77	17	312
d	237	20	73	125	924	466	171	2016
dl	140	12	52	77	466	410	100	1257
l	37	2	15	17	171	100	56	398
Totals	647	56	194	312	2016	1257	398	4880

SS. *Streatham. Sister Plants' Apical Flowers.*

First Sister Plant. Extent of Basal Patch.

Second Sister Plant. Extent of Basal Patch.	*n*	*ns*	*s*	*sd*	*d*	*dl*	*l*	Totals
n	186	53	433	128	226	47	186	1259
ns	53	66	263	83	140	25	46	676
s	433	263	1582	527	777	125	494	4201
sd	128	83	527	216	278	38	162	1432
d	226	140	777	278	386	83	292	2182
dl	47	25	125	38	83	18	83	419
l	186	46	494	162	292	83	270	1533
Totals	1259	676	4201	1432	2182	419	1533	11702

TT. *Kidderminster. Sister Plants' Apical Flowers.*

First Sister Plant. Extent of Basal Patch.

Second Sister Plant. Extent of Basal Patch.	*n*	*ns*	*s*	*sd*	*d*	*dl*	*l*	Totals
n	826	99	124	517	426	1	326	2319
ns	99	22	23	65	59	—	42	310
s	124	23	64	96	61	—	51	419
sd	517	65	96	680	494	6	255	2113
d	426	59	61	494	868	—	197	2105
dl	1	—	—	6	—	—	2	9
l	326	42	51	255	197	2	216	1089
Totals	2319	310	419	2113	2105	9	1089	8364

UU. *Parkstone. Sister Plants' Apical Flowers.*

First Sister Plant. Colour at Margin of Petal.

Second Sister Plant. Colour at Margin of Petal.	*r*	*rp*	*p*	*rw*	*pw*	*vw*	*w*	Totals
r	274	106	426	3	73	21	118	1021
rp	106	90	295	3	85	12	101	692
p	426	295	1860	7	423	150	646	3807
rw	3	3	7	—	1	—	3	17
pw	73	85	423	1	196	45	243	1066
vw	21	12	150	—	45	20	93	341
w	118	101	646	3	243	93	476	1680
Totals	1021	692	3807	17	1066	341	1680	8624

VV. *Crewe. Sister Plants' Apical Flowers.*

First Sister Plant. Form of Capsule.

Second Sister Plant. Form of Capsule.	Round	Conical	Flat	Totals
Round ...	314	258	155	727
Conical ...	258	342	179	779
Flat ...	155	179	92	426
Totals	727	779	426	1932

WW. *Parkstone. Sister Plants' Apical Flowers.*

First Sister Plant. Wrinkling of Petal.

Second Sister Plant. Wrinkling of Petal.	*f*	*w*	*sw*	*nw*	Totals
f	6	5	30	40	81
w	5	46	164	313	528
sw	30	164	1210	1661	3065
nw	40	313	1661	2906	4920
Totals	81	528	3065	4920	8594

XX. *Arncliffe. Sister Plants' Apical Flowers.*

First Sister Plant. Wrinkling of Petal.

Second Sister Plant. Wrinkling of Petal.	*f*	*w*	*sw*	*nw*	Totals
f	78	131	252	57	518
w	131	314	596	111	1152
sw	252	596	1310	296	2454
nw	57	111	296	260	724
Totals	518	1152	2454	724	4848

YY. *Streatham. Sister Plants' Apical Flowers.*

First Sister Plant. Wrinkling of Petals.

Second Sister Plant. Wrinkling of Petal.	*f*	*w*	*sw*	*nw*	Totals
f	14	69	113	149	345
w	69	410	742	936	2157
sw	113	742	1380	1704	3939
nw	149	936	1704	2472	5261
Totals	345	2157	3939	5261	11702

ZZ. *Kidderminster. Sister Plants' Apical Flowers.*

First Sister Plant. Wrinkling of Petals.

Second Sister Plant. Wrinkling of Petal.	*f*	*w*	*sw*	*nw*	Totals
f	2	28	59	17	106
w	28	444	1013	194	1679
sw	59	1013	3356	779	5207
nw	17	194	779	382	1372
Totals	106	1679	5207	1372	8364

ON THE INHERITANCE OF COAT-COLOUR IN CATTLE.

PART I. SHORTHORN CROSSES AND PURE SHORTHORNS.

By A. BARRINGTON and K. PEARSON.

CONTENTS.

		Page
(1)	Introductory	427
(2)	Material	430
(3)	Colour Inheritance in Blue-Grey Cattle	431
(4)	Horn Inheritance in Blue-Grey Cattle	435
(5)	Colour History of Shorthorns	436
(6)	Possible Results of Various Colour Matings and Mendelian Interpretations	440
(7)	Assortative Mating in the Shorthorn Population	444
(8)	On Direct Inheritance. First Degree	445
(9)	On Direct Inheritance. Second and Higher Degrees	448
(10)	On Collateral Inheritance. Whole and Half Siblings	453
(11)	General Conclusions	454

(1) *Introductory.*

As far as we are aware no statistical investigation into the inheritance of coat-colour in cattle has hitherto been published. An attempt will be made in the present memoir to cover some of this ground in a manner similar to that in which investigations have already been made in the case of the coat-colours of the thoroughbred horse* and of the greyhound†. It is, perhaps, idle to assert that we have throughout desired to give as complete and accurate an account as lay within our power of the mass of material available, and this without prejudice in favour of any theory. T[illegible]se acquainted with the spirit which animates the work of biometricians will scarcely doubt this; those who have throughout misrepresented the purport of mathematical investigations into heredity will doubtless continue to do so. At the same time it may not be without service to reiterate what has

* *Phil. Trans.* Vol. 195 A, pp. 92–102; *Biometrika*, Vol. II. pp. 229–240.

† *Biometrika*, Vol. III. pp. 245–298.

been previously stated as to the scope of biometric work in this field. Our object is to state first the proportion in which sub-classes exist in a "population," to measure next the degree in which these sub-classes are mated at random, and then to determine the extent to which these sub-classes reproduce themselves. We believe that a statistical record of this kind is of fundamental importance whatever physiological theory of heredity be adopted. In treating cattle we have done precisely what was done in the case of thoroughbred horses and greyhounds, namely, the latest available volumes of the studbook have been taken*, and the colour pedigree of individuals entered therein worked out. In the work on cattle this was done on standard pedigree slips, of which a sample is given on the following page. These were filled in where possible as far as the great-great-grandparents and these schedules, representing a random sample of the existing general population, form the basis of our tables. As in the case of the greyhounds, supplementary investigations were made by collecting special pedigrees as those of white individuals. These special pedigrees were *not* embraced in the general sample except in so far as they fell into the period there dealt with, but they served to illustrate special points which it seemed desirable to elucidate.

On the basis of the above random sample† we first determined the extent to which random mating occurs in the general population. To what extent is there a fashion or taste in mating cattle? We next enquired how far our tabulated results admitted of a Mendelian interpretation. Let us suppose this question answered in the positive or negative sense, as the case may be, it leaves the tabulated data absolutely where they were as to the further question we have to ask: What is the *average* degree of resemblance between relatives in a population? To what extent do the features of the ancestry impress themselves upon the offspring? This is not a matter of theory, it is simply a statistical enquiry as to the degree of independence of two distributions. It resolves itself mathematically into determining the degree of independence of probability in two "chance" distributions. In order to dispose of all doubts as to quantitative scale and as to "normal" distributions of characters, we have adopted throughout the contingency method of measuring the deviation from independent probability, and in particular that form of it spoken of as mean square contingency. This makes all our results for one race strictly comparable *inter se*. But the colour classifications being different for different species, it is not so legitimate to compare these results together. In order to throw light on this point other methods have been adopted occasionally for comparison and control; these will be referred to in their proper places.

* In the *second* paper on coat-colour in horses, we began with the earliest studbooks in order to test the influence of the far more frequent greys and blacks.

† The word "random" is used with reference to coat-colour. When one or both parents had appeared several times, no further offspring of such individuals were taken, in order to avoid overweighting certain individuals. This affected the cows more than the bulls, because the former are classified in herds with sometimes close inbreeding. Thus 1000 consecutive cows might possibly have given only 50 to 60 sires.

GENERAL HEREDITY SCHEDULE.

Species: Shorthorns. No. in Series, . Character: Coat-Colour. Recorder: A. Barrington.

Parents	Male.	Female	Grandparents	Male	Female	Gt Grandparents	Male	Female.
2 78665 Crusader	red		4 66864 Cornelius	roan		8 60073 Touchstone	r & w*	
						9 Butter Scotch		white
			5 Rosemary 93rd		red	10 62235 Bucaneer	red	
						11 Rosemary 90th		r & w
3 Reception 13th		white	6 68388 Commander	roan		12 58582 Cambridge Duke 26th	red	
						13 Lady May		roan
			7 Reception 12th		roan	14 65255 Cheque	r & w	
						15 Reception 8th		roan

Subject.	Male	Female	Siblings.	Male	Female.	First Cousins.	Type.*	Male.	Female.
v 49 p 888	^{0}a	^{1}a roan		^{0}b	^{1}b				
				^{0}c	^{1}c				
				^{0}d	^{1}d				
				^{0}e	^{1}e				
				^{0}f	^{1}f				

Remarks. Harforth Cymbal

Gt. Gt. Grandparents

52956	red
Sunflower	roan
58189	roan
Buttercup	r & w
57992	roan
Bella 6th	roan
57018	red
Rosemary 58th	roan
55632	red
Red Rose 13th	red
54870	roan
Lady Jane	r & w
62180	roan
Charity	r & w
60494	red
Reception 3rd	roan

* Father's Brother's Offspring $= 2_0$ Father's Sister's Offspring $= 2_1$ Mother's Brother's Offspring $= 3_0$ Mother's Sister's Offspring $= 3_1$

Any simple or complex Mendelian system, any theory of hereditary determinants, must lead when it is applied to a "population" to its appropriate system of contingency coefficients between relational grades. It may be easy or it may be difficult to work out their theoretical values, but they undoubtedly exist, as soon as we know the nature of the mating. And without knowledge of the amount of mating of sub-groups no determinantal system can be applied to wild populations. If proven a theory of determinants would immensely assist breeders, but it is extremely difficult to see how it could be applied to wild populations and to the Darwinian theory of evolution by inheritance and selection. Hence, if we lay special stress on finding directly the coefficients which fix the intensity of inheritance in populations, it is not because we should in the least undervalue the discovery of a well-demonstrated determinantal theory of inheritance, but because we should then have, before applying it to populations and seeing its full bearing on the problem of evolution, to actually go through the theoretical determination of constants, which can without any such hereditary theory at all be determined from the material at our disposal.

When, however, we have found the degree of resemblance between relatives quantitatively, that is to say, the empirical numbers which for any population mark the divergence from independent probability between any relational grades, we can next ask : Is there any connection between these numbers within the same species? and are the coefficients for different species, or for different characters in the same species alike? When we answer these questions we are determining whether heredity is the same or not for all characters and species, and to what extent ancestry and stock are of importance. It may be said that a valid determinantal theory would answer these questions also. So it undoubtedly would, but no such theory is, in our opinion, even approximately established as yet; and meanwhile there is and will remain every justification for attacking the problem by a method, which starts from no hereditary hypothesis, but merely endeavours to interpret the observed numbers of each class.

So much may be said of the methods adopted in this paper. It is a repetition of what has already been written in this *Journal*, but it is desirable to restate it here in order to emphasize the aim of the present investigation.

(2) *Material.*

When we determined some years ago to consider the inheritance of coat-colour in cattle we directly sought for a simple case which might possibly enable us to convince ourselves of the efficiency of a quite brief Mendelian formula. The Guernsey cattle have a very considerable variety of coat-colour. With shorthorns a briefer classification is sufficient, but still it is wide enough to allow of the possibility of very complex allelomorphs. We were aware that the white shorthorn breeds nearly true, and that the black Galloway does the same. Hence the cross of the white shorthorn and the black Galloway seemed to offer a possibility of testing the simplest Mendelian formula. Unfortunately the resulting blue-grey

cattle, although of great value for trade purposes, do not form the subject of any herdbook record. Still it seemed possible to obtain some simple results by an appeal to large breeders of blue-greys. Their returns form the subject of the first part of the present paper. We next endeavoured to obtain some general idea of the colour inheritance of British cattle, but the colours of the original progenitors of existing types seem very hypothetical, and our researches in this direction, embodied in a later part of this paper, are not very helpful. Attention was then devoted to two types: the shorthorn and the Guernsey. In the case of the former *Coates' Shorthorn Herdbook*, issued by the Shorthorn Society of Great Britain and Ireland, has now reached the fiftieth volume, and we have to heartily thank the Society for placing a copy of it at our disposal for the present investigation. This herdbook is very complete in character and deals with very large numbers of beasts. At least some breeders insert the coat-colour of calves sent to the butcher, but we doubt whether such entries are universal. In the case of the Guernseys, the *English Guernsey Herdbook* has appeared for eighteen years, the pedigrees lead fairly soon back to the *Guernsey Guernsey Studbook*, of which ten volumes had appeared up to date. We have to heartily thank the English Guernsey Society for placing a copy of their herdbook at our disposal. We supplemented it by the purchase of the *Guernsey Guernsey Studbook*. The latter seems to us a less reliable work, as cattle are included with many coat-colours, which would not be admitted now into the English Guernsey record. But a full discussion of this matter is postponed until we come to deal with Guernseys in the second half of this paper. Pedigrees with the statement of coat-colour were filled in for shorthorn and Guernsey cattle from these herdbooks, and to obtain an adequate supply of such pedigrees has demanded a large amount of very laborious work. We shall now take our subdivisions in order:

(3) *Blue-Grey Cattle. Colour Inheritance.*

In this case the blue-grey cattle are as a rule produced by crossing a black polled Galloway cow with a white shorthorn bull. The white shorthorn (with black eyes) and the black cattle breed relatively true, so that if we term one B and the other W we might test a simple Mendelian formula $(BB) \times (WW) = (BW)$, and then further investigate the crosses $(BW) \times (WW)$, $(BW) \times (BW)$ and $(BW) \times (BB)$.

As in the case of Japanese waltzing mice crossed by Albino white mice we do not meet with any simple phenomenon of dominance, the general coat-colour of the offspring being blue-grey. The problem therefore turns on the action of these hybrids. Our first enquiry was as to the nature of the shorthorn bull purchased by the blue-grey cattle breeders. Mr H. L. Fife, of the Raby Estates Office, speaking of experience with the sales of produce from the Raby herd, writes:

"Breeders of blue-grey cattle in the North-West of England, who produce the same by crossing a white shorthorn bull with a black polled cow, are always most particular when buying a white sire to satisfy themselves if possible, that such sire is the produce of two white parents. If he is not, that is to say, if the parent on one side only is white, their theory is that his stock when crossed with black cows do not get the blue-grey colour they like."

This statement led to the framing of the following questions, which were issued in a circular form to the breeders of blue-grey cattle*:

We are told that breeders of blue-grey cattle object to white shorthorn bulls, whose parents were not both pure white as giving bad results.

(i) Can you from your own experience tell us whether you have bred from such a shorthorn bull, and if so whether you found it to give bad colour results and what these results were?

(ii) If a blue-grey cow be crossed with a white shorthorn what is the resulting coat-colour? Could you give us actual instances of this cross? In particular do pure white cattle ever result from this cross?

(iii) If the blue-grey cross be covered by a Galloway bull, what is the resulting coat-colour? Do black as well as blue-grey offspring result? Can you give us instances in your own experience?

(iv) If the blue-grey cross be mated with itself, blue-grey cow with blue-grey bull, what is the resulting coat-colour? Do you know instances of black cattle or pure white cattle resulting from this cross? Are the majority of the offspring blue-grey?

It will be observed that questions (ii), (iii), and (iv) deal respectively with the crosses $(BW) \times (WW)$, $(BW) \times (BB)$ and $(BW) \times (BW)$ and were intended to elicit information with regard to these matings. Question (i) was framed so as to draw forth either facts or impressions of the breeder as to the need for purity in the white ancestry.

These questions produced a number of answers—the experience of all breeders being by no means alike—which we have summarised below. We have most heartily to thank these gentlemen for their kindness and invariable patience in replying to our queries.

There appears to be general assent to the proposition that a pure bred Galloway cow and a white shorthorn bull of white parentage give the desired blue-grey cross.

When the white shorthorn bull is *not* from white parentage, Messrs Hyslop state that besides greys, there appear reds, blacks, particolour black and white, and occasionally a white with dark nose and insides of ears. Mr de Vere Irving prefers to use a shorthorn white bull of white parentage on both sides. He finds on using a white bull of roan parentage, that the offspring are liable to come dark, some almost black in colour, and others very dark blue-grey. Mr G. Hodgson always tries if possible to get a bull from white parents, which he finds makes the progeny the desired *light* blue-grey. If a white bull of white parentage of a satisfactory type is not to be had, then he takes one from roan parents, but this in his experience makes the colour of a darker nature. Mr R. Tinniswood has bred from a pure bred Galloway and a white shorthorn bull not of white parentage and has had occasional coloured calves from this cross. He considers that it is very desirable that the white shorthorn bull should be from white parents to get the best results. Bulls not from white parents mated with pure bred Galloways in the majority of cases give the right colour, but may give reds, roans† or nearly blacks. Mr Tinnis-

* We have to thank Messrs Harrison and Son of Carlisle for kindly providing us with a list of blue-grey cattle breeders.

† The red and roan hair we have examined is exactly alike in appearance to that of red and roan shorthorns.

wood considers that the ancestry of the Galloway cow is even more important than that of the shorthorn bull; and states that the *true pedigree* black Galloway crossed with the white shorthorn of non-white parentage will give a blue-grey, though some of the offspring will be very dark blue-grey. The whole of this evidence seems to show that black Galloway × white shorthorn* does not give invariably the desired light blue-grey blend, but that the deviations from this colour in the opinions of experienced breeders depend on the ancestry of the (BB) and the (WW). The hybrid generation is not uniform unless the parents come of selected stock. Such appears to be the best answer we can get to our first problem.

Our second question relates to the crossing of the hybrid (BW), the blue-grey cattle with (WW) the white shorthorn. This ought to give half blue-grey (BW) and half white shorthorns (WW). The experience of breeders is not uniform, but most probably it is based on white shorthorn bulls with varying colour pedigree.

Mr Tinniswood finds that the cross is invariably a white calf with a black nose and ear linings. He has been careful to use shorthorn bulls which as to ear and nose are perfectly white. Mr Hodgson has frequently bred this cross and finds the results always white, occasionally with a black nose. He says this cross gives white cattle which are not, however, to be reckoned as pure white shorthorns. Thus in the experience of both these breeders (BW) has not occurred as a result of $(BW) \times (WW)$, nor in the experience of at least one of them (WW), but white with black markings. Other breeders, however, have had different experiences. Mr G. Richardson finds all colours to arise from the cross. Messrs Hyslop find like Messrs Tinniswood and Hodgson that white with coloured nose and ears was the resultant coat-colour, but some light blue-grey, some "grizzled" (i.e. red mixed with white hairs), and some black generally at base of coat. Mr de Vere Irving considers the coat-colour uncertain, and cites as illustrations, a blue-grey cow × white shorthorn bull giving a white calf with black nose and blue-grey ears. Another blue-grey cow with the same bull giving last year a blue-grey calf and this year a roan. He has also had *pure* white cattle from this cross. We think we may conclude from these results that the essential feature to which Mendel drew attention, i.e. segregation in the second crosses, certainly does occur, but its nature is not compatible with any simple Mendelian formula. We have not only (WW) and (BW) occurring from $(BW) \times (WW)$, but also red and roan appearing in the hair, as well as white with black markings which did not exist in either the pure bred Galloway or the shorthorn parents. Further, this segregation although possibly more marked than in the hybrid generation is to be found in the hybrid generation itself, if the cow and bull while presenting the external features of pure Galloway and pure shorthorn, have not pure black and pure white ancestry.

We now turn to the cross $(BW) \times (BB)$. This should give (BB) and (BW) in equal proportions. Mr de Vere Irving finds the resulting offspring black or *dark*

* The usual cross is $♀ (BB) \times ♂ (WW)$, but the reciprocal cross is sometimes made. The produce of this cross is said, however, not to be uniform in colour. It is adopted because shorthorn heifers can be bought on better terms than pure bred Galloway heifers; while the shorthorn heifer as a cow makes a better price than a Galloway for towns' dairying.

blue-grey. He has at present a black heifer from this cross and has personally known instances of the dark blue-greys. Mr Richardson's experience is the same as Mr Irving's, i.e. black and blue-grey. Mr Tinniswood's experience is that the offspring take black, but occasionally, especially if from *light* blue-grey cows, the black is accompanied by white markings. Mr Hodgson finds the coat-colour not very certain, generally black; but from a *light* blue-grey cow he has had blue-greys and roans; from a *dark* blue-grey cow, generally black with only an occasional roan. This is interesting because we have seen that the *light* blue-grey results from the purity of the Galloway cow and the white parentage of the shorthorn bull. It would thus appear that the grandparental ancestry influences the nature of the cross $(BW) \times (BB)$, through the character of the (BW) itself. Messrs Hyslop have had still more varied results, namely, dark blue-grey, some "grizzled" (red and white hairs), black, and white with black nose and ears. In other words if white with black nose and ears be taken to represent the (WW) arising from $(WW) \times (BW)$, it cannot represent the (BW) in $(BB) \times (BW)$.

We now turn to the last problem: what does $(BW) \times (BW)$ give? Messrs Tinniswood and de Vere Irving have never tried the cross. In fact it is clearly not one that blue-grey cattle breeders are likely to make*. Mr Hodgson has had no personal experience of this cross, but has seen instances in which a blue-grey bull being used on blue-grey cows, blue-grey and roan offspring resulted. Mr Richardson has known instances in which this cross being made the offspring were nearly all blue-grey, but the following season with the same cows and bull, red, roan, white, black, with only a few blue-grey calves appeared. Messrs Hyslop state that the cross gives blue-grey, black, red, "grizzled" (red with white hairs) and an occasional white. They know no breeder, however, who makes at present this cross although it used to be followed. So far as this experience goes therefore $(BW) \times (BW)$ appears to give (BW), (BB) and possibly (WW), although it is not white with black markings, such as occurs in the $(BW) \times (WW)$ and $(BW) \times (BB)$ crosses. But on the other hand red hair has appeared when the original pure bred Galloway and shorthorn did not patently contain this colour.

To sum up, we may conclude, we think:

(*a*) That when pure black and pure white parents of races which breed nearly true are crossed, the hybrids are not homogeneous, there is a certain amount of segregation even in this generation, and colour may appear which was latent in the parents. It seems to us a difference in words only, whether we attribute this appearance to a latent determinant, or say that it is reversion to an ancestor. The pedigree is, in fact, the only *à priori* guide the breeder has to the possibility

* Mr Tinniswood breeds about 100 blue-grey calves a year. As he remarks, it is possible to get blue-greys from crossing a blue-grey cow with either a Galloway or shorthorn bull, but the colour "very seldom occurs." Hence the breeder naturally turns to the white shorthorn and Galloway first cross. Mr Tinniswood has used both crosses, generally the shorthorn bull, but also the *pure* shorthorn white cow and the Galloway bull, a cross he has never known to give red or roan. He has known red or roan from the reciprocal cross, and attributes it to want of pedigree in the Galloway cow.

of such reversion. As he puts it, "it is desirable that the white shorthorn should be of white parentage," and "the Galloway a pure Galloway."

(*b*) More segregation takes place when the hybrid generation is crossed with either original stock. Red hair appears, the white shorthorn is replaced by white with black nose and ears, and this or pure white appears from the cross $(BW) \times (BB)$. Our hope therefore that a simple Mendelian formula might be tested on this case—the only one in which we knew of two whole colour cattle of different races being habitually crossed—has failed. Our enquiries were, we believe, justified, however, by the apparent simplicity of black × white = blue-grey*. We did not think actual colour statistics in this case worth collecting, for if we must deal with more complex colour categories, it is at once best to turn to breeds like the shorthorn and Guernsey, for which ample colour statistics are already on record. So far as it is possible to make a comparison between rough experience of the present kind on blue-grey cattle and more definite quantitative experience on mice, the colour results of crossing are in many points strikingly like the Oxford experience with Japanese waltzing and albino white mice. It will need a complex allelomorph to describe these colour changes, if indeed the phenomena can be thus described at all.

(4) *Blue-Grey Cattle. Horn Inheritance.*

It seemed worth considering whether a second character, the horns, would admit of a simple Mendelian analysis. The black Galloways are polled cattle†, and the white shorthorns, horned. We can represent these respectively by (PP) and (HH), and we have to enquire as to the following results: $(PP) \times (HH) = (PH)$, $(PH) \times (PP)$, $(PH) \times (HH)$ and $(PH) \times (PH)$.

First, $(PP) \times (HH)$ gives theoretically the blue-grey (PH). Actually Messrs Harrison say (PH) are "chiefly" polled, and Mr de Vere Irving finds that about one in twenty of (PH) is horned. Thus "polled" is dominant, but it is not absolutely dominant. Mr Irving has not found that exceptions are to be associated with want of pedigree in the Galloway, and he has found no difference with regard to horns between ♂ $(PP) \times$ ♀ (HH) and the more usual cross ♀ $(PP) \times$ ♂ (HH). Mr Robert Tinniswood agrees on this point, and finds the result polled if both sire and dam are pure bred. Mr Hodgson finds not more than about 1 in 30 to 40 horned; he considers that this is usually due to impurity in the Galloway pedigree, but horns may occur with exceptionally well bred sire and dam.

$(PH) \times (PP)$. This cross should give (PP). Mr Irving finds the offspring are polled. Mr Richardson says that they are all polled. Mr Tinniswood has not

* As a matter of fact this blue-grey is, to judge by hair samples, simply a blend of white and black hairs, varying in proportion from almost pure black to a large majority of white.

† Our English informants state that the Galloway is in their experience polled. If the breed now contained (PH)'s, (HH)'s would appear. Boyd Dawkins, *Quart. Journal Geol. Soc.* 1867, p. 177, says that in the 18th century Galloways were not polled, but his evidence for universality of horns seems far from conclusive. Further details as to the crossing of polled and horned breeds will be given in the next number of this Journal.

largely bred this cross but all the calves have been polled. Mr Hodgson on the other hand has found 1 out of 8 to 10 horned.

$(PH) \times (HH)$. This should equal $2(PH) + 2(HH)$ and give half horned and half polled. Messrs Hyslop find the majority horned. Mr Richardson finds three-quarters horned, and Mr Irving that the horned are in a majority of three to one. Mr Irving says that it sometimes happens that the horns do not grow to much size but in the majority of cases they are there, and he does not consider that such cattle can be described as polled. Mr Tinniswood finds the calves bred this way generally horned, perhaps 2 or 3 in twenty would come polled. Mr Hodgson on the other hand finds in this case only 1 in 4 to 6 horned and illustrates by three blue-grey ♀s put to a white shorthorn bull last year giving all polled calves.

$(PH) \times (PH)$. This should give $(PP) + 2(PH) + (HH)$, which being interpreted is polled in a majority of three to one. Few breeders make this cross, and Messrs Irving, Richardson and Tinniswood have had no experience of it. Mr R. Hyslop finds six polled to one horned, and Mr Hodgson says one out of 6 to 8 horned. These numbers are far from the Mendelian ratio.

Thus, as in other cases which it has been our misfortune to come across, we find approximations to Mendelism, but we do not find Mendelism an accurate account of the records provided. The polled (PH) might really be a (PP), it cannot be an (HH), and therefore we could account for less than a half being horned in the cross $(PH) \times (HH)$, but not for more than a half being horned. The only explanation could be that polled in a number of cases are latently horned, i.e. we cannot distinguish (PP) from (PH). This would account for the horned cattle appearing in $(PP) \times (HH)$, but not for the usual absence of horns in $(PH) \times (PP)$, for some reputed (PP)'s would be (PH)'s. Anyhow the further pursuit of a somewhat elusive Mendelism in blue-grey cattle did not seem to us likely to be profitable.

(5) *Colour History of the Shorthorn.*

The classification of shorthorns into broad categories is extremely simple as compared with some other types of domestic cattle. The recognised classes are *Red* (R), this is supposed to contain whole reds; *Red with a little White* (RLW) or red with occasional white spots or with white markings; *Red and White* (RW), i.e. particolour; *Roan* (Ro), a mixture of red and white hairs; and *White* (W). The white are not albinos*, having black eyes and occasionally black noses. Attempts more or less successful to get rid of the black noses have been energetically made by breeders. The existence of black pigment in the nose is probably not always recorded in the herdbook records. We have seen that breeders of blue-grey cattle strive to get shorthorn bulls without black noses, but as there is always black pigment in the eye, the question of absence of black pigment in the nose is perhaps not of first-class importance in the problem of crossing. It appears perfectly certain that eye and nose pigmentation in intercrossing white shorthorns

* A herd of true albino white cattle (whole white and pink eyes) has been reported to us as existing in Sweden, but we have not succeeded in locating it.

does not now introduce black into the coat-colour, though how far it may appear as red may be open to discussion. Red in the shorthorn appears in different shades roughly classified by breeders as dark and light, and we have corresponding dark and light roan, according to the character of the red hair in the roan. These finer distinctions are not recorded in the herdbook*.

From the standpoint of our present investigation, it would be of very great suggestiveness if we could determine in number and colour the original factors of the modern shorthorn. So far as we have been able to gather from writings on the subject there seems to be a good deal of diversity of opinion among authorities as to the origin of the domestic breeds of British cattle. There are three somewhat divergent main views. First that of Boyd Dawkins†. He identifies the *Bos Urus* of the prehistoric and historic periods with the *Bos primigenius* of the Pleistocene period, but he does not consider that either survived in this country as source of the wildpark cattle. He holds that the small dark Welsh and Scotch mountain cattle represent an original Celtic shorthorn, and that the Romans used this Celtic cattle and did not import larger breeds of their own. This whole coloured Celtic shorthorn was driven out not by the Romans, but by the Anglo-Saxons. The Chillingham wild cattle on this view are living representatives of the breed brought by the Anglo-Saxons. They are not survivors of the *Bos Urus* in this country, but imported Anglo-Saxon cattle run wild‡. At first sight there does not appear the same reason for the disappearance of the Anglo-Saxon cattle, that there was for the expulsion of the Celtic cattle, and Boyd Dawkins' citations from Welsh Laws may be read in more ways than one, for fines have generally become customary many generations before they are codified. It is also unlikely that all the place-names in Ireland referring to white cattle are later than the Saxon invasion.

Another view has been taken by McKenny Hughes§. He considers that the characteristics of the *Urus* nowhere appear in the Romano-British cattle. The Kerry cattle are typical of the Celtic shorthorn, the Chillingham cattle are the nearest representatives of the breed introduced by the Romans‖. The Highland and Welsh cattle are derived largely from the Roman breed with admixture of the Celtic shorthorn. These are all whole coloured or shaded. The *Longhorns*, frequently particoloured or sheeted, are the offspring of breeds introduced from

* The range and continuity of the reds seem fairly obvious from the samples of hair colour in our possession.

† *Quarterly Journal of the Geological Society*, Vol. xxii. 1866, pp. 391–401. "Fossil British Oxen," *Pleistocene Mammalia, Palaeontographical Society*, Boyd Dawkins and Sandford, 1866–1886, in parts.

‡ B. A. Report on Herds of Wild Cattle, *Report*, 1887, p. 135, appears to leave the question of importation open. Rütimayer ("Ueber Art und Race des zahmen europäischen Rindes," *Archiv für Anthropologie*, Bd. i. S. 219–250) holds that the English wild cattle, the large horned Hungarian and *Italian* cattle belong to the Primigenius race.

§ *Archaeologia*, Vol. lv. Part i. pp. 125–158, 1896. "On the more Important Breeds of Cattle which have been recognised in the British Isles, etc."

‖ A comparison of the skulls of Chillingham and Italian bulls with that of the *Bos primigenius* can be easily made at the Natural History Museum, South Kensington.

the Low Countries and Holstein in later mediaeval times. McKenny Hughes says that the Durham and modern shorthorns belong to a much later period and fall outside his enquiry. Thus writings on the subject do not give us very definite results for our present purposes. We further appealed to Professor James Wilson, of the Royal College of Science, Dublin, who is an authority on the matter, and who most kindly provided us with the following brief statement:

(*a*) The first cattle in Britain were of a large type, *Bos primigenius*, which became extinct here in the stone age, although possibly surviving later on the Continent. If they had anything to do with existing British cattle, which is extremely doubtful, the relationship is very remote and is only to be traced through later importations from the Continent.

(*b*) The next type to be noted in Britain is the Celtic shorthorn, *Bos longifrons*, a much smaller animal. Patches of hair have been found in the caves, and these are said to be of a dark colour. Professor Wilson is strongly of opinion that *B. longifrons* was mainly black. It has been said to be red or black.

(*c*) The next invasion, that of the Romans, brought the Roman white cattle, which probably had black markings. The Romans formed joint settlements with the British, and a Romano-British cattle resulted in the neighbourhood of Roman settlements, from crossing the Roman and Celtic types.

(*d*) In the next place we have a Saxon invasion, bringing with it Saxon cattle. There was probably some mixture of breeds on the Celtic and Saxon frontiers, but generally speaking Saxon cattle occupied the parts from which the Romano-British population had been expelled. These cattle were red.

(*e*) Following on the Anglo-Saxon was a Dutch invasion of cattle into the east of England. By Dutch we understand from the "Low Countries." This invasion began in the middle ages and continued till a century ago, and again involved a certain amount of mixture. The Dutch cattle were probably "flecked," the flecking being white with black, red, or brown. The Englishman developing a prejudice against black, we may suppose the reddish mixture to remain, and to have become the basis of roan.

Before, therefore, the beginning of pedigree cattle-breeding, or earlier in the sixteenth century say, we should expect a distribution of cattle races much as follows:

Wales and Scotland, Celtic or black cattle, with some of Roman descent and perhaps some whole reds.

Midland England, very mixed cattle, compounded of the following races in order: Celtic, Roman, Saxon and Dutch. The colours were black, white with black or brown points, red and flecked. Round the edges of the country the breeds were comparatively pure, but in the Midlands, where they all met, they were a mixture; the Romano-Celtic element probably prevailing.

Southern England, Anglo-Saxon, red.

Eastern England, Dutch cattle, flecked or roan.

The purest British cattle would thus be found in Wales, Scotland (and West Ireland), where black is the prevailing colour.

The Romano-British *were* up the west and middle of England, where they still survived to about a century ago. They were the old longhorn breeds.

The wild white cattle of Chillingham, Chartley, and elsewhere, are Roman cattle run wild owing to the unrest of the Anglo-Saxon, and later and more especially the Danish invasions. Traces of the Roman cross are still to be seen among Scottish, Welsh, and Irish cattle. The wild white cattle have black muzzles and black or brown points; that is their hair is brown, or dark at the ears, round the eyes and muzzle and often at and below the knees.

The Anglo-Saxon red cattle, which were once spread over the south of England, are still to be found in the red cattle of Norfolk, Sussex, Devon, and Hereford.

The Romano-British or longhorns were some black, some white, but mainly flecked, and were largely driven out by the flecked Dutch, although some mixture probably took place.

It will be seen that Professor Wilson's account differs to some extent from both those of Boyd Dawkins and McKenny Hughes.

Upon the ingredients just referred to the breeders had to work when pedigree cattle-breeding, which is scarcely more than a century old, came into vogue. The shorthorn has possibly arisen from four races, the Celtic, a Romano-British, an Anglo-Saxon, and the "Dutch*", and even some of these are mixtures.

Confining our attention to coat-colour, the black Celtic, if it contributed anything, must have been bred out early. But it is doubtful whether there was not a red Celtic, and a red Galloway heifer (1760–1770) appears to have been connected with the beginnings of the breed. Thus the shorthorn red may have had three sources, an Anglo-Saxon red, the red of the Dutch flecking, and the supposed Celtic red. The white may have come through the Romano-British, through an Anglo-Saxon white or possibly through the white in the Dutch. The particolours and the roans are of equally doubtful origin, although it probably is safe to assert that they are due to the breeds of latest importation; and it thus seems fairly impossible to determine *à priori* how many distinct red, roan, particolour, or white types may really exist in the case of the shorthorn. The importance of this statement for any Mendelian interpretation must be obvious. We may have reds which are dominant, recessive or even heterozygous to white or even to other reds, and the search for a Mendelian formula becomes very elusive.

If we turn from the possible ingredients of the shorthorn breed to the history of its origin we find matters still more difficult to disentangle.

There appears, by the will of John Percy of Harum, to have been a breed of shorthorns in Yorkshire as early as 1400, and these were not whole colour. In

* Their nearest Continental cousins to-day are the German *Fleck-Vieh*.

1640 we again hear of a shorthorn breed; and a native race with short horns, large bodies, and black or red in colour, existed in Holderness and on the banks of the Tees in the middle of the eighteenth century. A new breed of shorthorns was introduced from Holland in the eighteenth century, and much improvement was attributed to the introduction of Dutch bulls, but it is said that these were the offspring of English shorthorn cows sent by James II to William of Orange. In the early shorthorn pedigrees we find some black, flecked, and spotted beasts. However debateable the origin of the shorthorn may be, it appears to have existed as a distinct breed by the middle of the eighteenth century, and possibly 50 or 100 years previously. Its source is probably Yorkshire and Durham with Dutch reinforcements. The early breeders seem to have had much the same colours, red, roan, white, and their mixtures. One noteworthy early breeder introduced a cross which is supposed to have become a potent factor with a *red* Galloway polled cow. The blacks and fleckeds rapidly disappeared. The proportion of colours was, however, somewhat different. Thus from pedigrees before 1850 we find:

Red	R. W.	Roans	Whites	
14·5	42	38	5·5	per cent.

and from recent statistics:

Red	R. L. W.	R. W.	Roan	White	
30	6·5	4·5	54	5	per cent.

Thus the whites have remained about stationary in number, but the whole reds and roans have increased at the expense of the particoloured, and this even if we throw the red with white markings into the particolour.

It cannot be said that this brief examination* casts much light on the possible components of the modern shorthorn. We do not know the actual constituents of the early Yorkshire shorthorn. It suggests that the white may have come from Holland as a whole colour and not be a contribution of Saxon or Romano-British factors. We see a definite Dutch contribution and a possible Celtic red through Collings' Galloway cross. We do not know any more certainly whether an Anglo-Saxon red also contributed. Failing this, we are thrown back on actual breeding experiments to disentangle possible Mendelian factors.

(6) *Possible results of the various colour matings and Mendelian interpretations.*

There are fifteen colour crossings possible with our five shorthorn colour categories, and since each crossing might give offspring of one of the five categories, 75 possibilities arise.

* See: *Thomas Bates and the Kirklevington Shorthorns*, by C. J. Bates, 1897. Bell's *History of Improved Shorthorn Cattle*, 1871. B. Rudd: *An Account of some of the Stock of Shorthorned Cattle of Ch. and R. Collings*, 1821; *Origin and Pedigree of the Sockburn Shorthorns*, 1822. Lewis F. Allen: *The American Herdbook, to which is prefixed a Concise History of English and American Shorthorns*, 1856, and *History of American Cattle*, 1868.

These categories are arranged in the following table:

TABLE I.

Mating	Possible Offspring				
	R.	R. L. W.	R. W.	Ro.	W.
R. × R.	×	×	×	×	*wp*
R. × R. L. W. ...	×	×	×	×	?
R. × R. W.	×	×	×	×	*wp*
R. × Ro.	×	×	×	×	×
R. × W.	×	×	×	×	*wp*
R. L. W. × R. L. W.	×	×	×	×	*wp*
R. L. W. × R. W. ...	×	×	×	×	*wp*
R. L. W. × Ro. ...	×	×	×	×	*wp*
R. L. W. × W. ...	*wp*	?	*wp*	×	*wp*
R. W. × R. W. ...	×	×	×	×	*AHB*
R. W. × Ro. ...	×	×	×	×	*wp*
R. W. × W. ...	×	*wp*	×	×	*wp*
Ro. × Ro.	×	×	×	×	×
Ro. × W.	×	×	×	×	×
W. × W.	*wp*	?	*AHB*	*wp*	×

All the categories marked × were found in the first random sample of 2172 calves taken out of the Herdbook. Some of course are infrequent, but their relative frequency will be found at once by examining Tables II and III. Those marked *wp* came to light in following up the pedigrees of white cattle. They are taken from vols. 37 to 49, in the bulk of cases from vols. 47 to 49; they are not exhaustive, but in most cases several schedules with each type occur in our record. The two cases marked *AHB* were noted in a single volume of the *American Herdbook*. No stress is laid on them.

Thus out of the 75 possible arrangements all but three have been found, and these three are all concerned with the comparatively scarce category of red cattle with white markings. Many of the 75 groups are rare, some excessively rare. It is always possible to assert that the rarest of them are due to misprints or to breeders' mistakes. We have made enquiries of breeders and in some cases additional instances have been provided, but we have kept to the cases actually on Herdbook record*.

Take as an illustrative case $W \times W$. This is a comparatively rare cross; because some breeders think white cattle delicate, others do not breed it because the produce have less sale and at lower prices in the export market. It is either unfashionable or unprofitable. Most breeders assert that $W \times W$ always gives W; it *generally* does, but their experience when you come to actual numbers must be very limited. In fourteen volumes of the Herdbook containing about 57,000 calves, 91 cases only of $W \times W$ were found, though quite possibly

* Mr T. Milne tells us, for example, of an all red heifer turned down with two young bulls, one all red and the other *RLW*, producing from one or other a white calf.

some few were overlooked. So that the cross is one of extreme rarity. In these 91 cases the calves were in 86 cases white, in four cases roan*, and in one case red†. The *American Herdbook* examined for other matters produced incidentally a case of $W \times W = RW$. Thus *Coates' Herdbook* indicates that in 5½ per cent. of cases white does not breed true to white. We are therefore driven to the conclusion that, if the records are to be trusted, there are latent determinants in white shorthorns or in some white shorthorns which under certain circumstances can be called into play. This view is confirmed by another consideration. Some wild white cattle, like the late Chartley Herd with its representatives at the Zoological Gardens, have black muzzles and black ear-linings. They are reputed to drop occasionally black or red calves. Other herds of wild white cattle have black noses and red inside the ears, and have bred absolutely true to this type for at least a century. As we have seen, the white shorthorns breed generally true to their type. The cross wild white cow × shorthorn white bull gives a hybrid pure white cow, but such pure white cows, if crossed with a shorthorn white bull, may produce a considerable percentage of both roan and red as well as white calves. In other words, while wild white cattle of the type considered breed absolutely and the white shorthorns sensibly true to type, and while the hybrid is pure white, the hybrids' offspring by white shorthorns segregate into colours occurring in the near ancestry of the shorthorn, and possibly but not probably in the distant ancestry of the wild white. These facts, coupled with the experience that white shorthorns if crossed will, although rarely, give roans and reds seem to show the existence of latent colour determinants in the whites depending upon their ancestral history. The phenomena described are similar to those noted in albino white mice, which carry in latent form traces of their colour ancestry, traces which can become patent under various crossings.

The simplest Mendelian formula would be based on considering the white as recessive and the red as dominant, and the roan as hybrid, i.e.

$$(RR) \times (WW) = (Ro),$$

and this is *roughly* an expression of some of the facts of the case. Thus $(WW) \times (WW)$ gives (WW) in all but 5½ per cent. of cases. In 196 crosses of $(RR) \times (RR)$ we find 156 give (RR), or in about 79·5 per cent. of cases. But, besides about 3 per cent. of roans, we have 12 per cent. of calves with white markings and 5·5 per cent. of particoloured. We are thus compelled to consider all red, as we have just considered all white, as either consisting of several subclasses only to be discovered by breeding, some all reds being really heterozygous, or else as containing latent white marking and particolour determinants. But the difficulty does not stop here, there appear to be authentic cases of $(RR) \times (RR)$ giving (WW). Thus we find Favourite Duke (R) got out of Clara 3rd (R) not only a normal red bull calf, but a red and white cow calf and the *white* cow calf Clara 5th‡ and the (R) Heydon

* Bull calves: Unnamed, Vol. xxxvi. p. 277; Shaftesbury, Vol. xxxvii. p. 292. Cow calves: Crystal Valley, Vol. xliv. p. 273; Seam, Vol. xlviii. p. 809. Other cases in *American Herdbook*.

† Cow calf, Bright Anna 2nd, Vol. xliii. p. 318.

‡ *Coates' Herdbook*, Vol. xli. p. 623.

Rose crossed by Earl Clarence 2nd (R) gave the *white* cow calf Minnette*. Such cases may be very rare indeed but, if authentic, reduce the Mendelian formula to a rough empirical statement of a statistical ratio; they are inconsistent with any theory of a pure gamete.

If we pass from the question of purity in the supposed homozygous elements to the first hybrid generation, we find in the Tables II. and III. of our random samples 90 cases of $(RR)\times(WW)$, and the result is (Ro) the assumed hybrid in 85 cases or in 94·5 per cent. of cases, but particoloured (RW) has occurred in two cases, white marking in one case and whole *red* in a third; while travelling beyond the field of our working sample $(RR)\times(WW)$ is actually recorded to have given (WW). Crossing the hybrid generation together, we should expect:

$$(Ro)\times(Ro)=(RR)+2(Ro)+(WW).$$

Our actual numbers on the random samples are:

$$514\,(Ro)\times(Ro)=86\,(RR)\dagger+31\,(RLW)+35\,(RW)+278\,(Ro)+84\,(WW),$$

as against the Mendelian formula:

$$=128{\cdot}5\,(RR)\qquad\qquad+257\,(Ro)+128{\cdot}5\,(WW).$$

The (Ro) is thus seen to be, like (RR) and (WW), a non-homogeneous group; we are compelled to suppose that some (Ro) have white marking and others particolour determinants latent in them. If we cross the hybrid (Ro) by the original stocks, we have:

$$656\,(Ro)\times(RR)=243\,(RR)+46\,(RLW)+39\,(RW)+324\,(Ro)+4\,(WW),$$

as against the Mendelian formula:

$$=328\,(RR)\qquad\qquad+328\,(Ro).$$

The close approximation to the Mendelian number of (Ro) is noteworthy, but the appearance of $4\,(WW)$ is again impossible unless some of the reds are to be treated as heterozygous. The latent white marking and particolour determinants are again manifestly needed.

Lastly, we take the case of $(Ro)\times(WW)$:

$$74\,(Ro)\times(WW)=1\,(RR)+1\,(RLW)+1\,(RW)+47\,(Ro)+24\,(WW),$$

as against the Mendelian formula:

$$=\qquad\qquad\qquad\qquad 37\,(Ro)+37\,(WW).$$

Here again we are compelled to assert a heterozygous white, or if we treat the individual instances of (RR), (RLW), and (RW) as heterozygotes and club them with the (Ro), the proportion 50 to 24 is very far from the Mendelian equality. Of course the same process could not then be applied to explain the anomalies in the crosses $(Ro)\times(Ro)$ and $(Ro)\times(RR)$. The anomalous cases in these crosses can

* *Coates' Herdbook*, Vol. xxxviii. p. 231.

† No attempt is made to express anything but (RR) and (WW) in Mendelian form, the constitution of the other individuals being unknown.

be multiplied by seeking for them and, not doing as we are here doing, dealing with the material provided in the pedigrees of a random sample from a few volumes of the *Herdbook*. It would thus seem that no simple Mendelian formula can possibly fit the shorthorn cases. Roughly, such a formula approaches the data in one or two points, but the roughness appears inconsistent with a theory of Mendelism being due to the purity of gametes. It is of course clear that the introduction of a complex allelomorph may improve matters, or the differentiation of whites and reds into different classes, homozygous and heterozygous. Increase in the number of available variables usually does give better fits. But the attempt to differentiate cows by the results of breeding from them must shatter on the paucity of their offspring. A cow will frequently give no more than 8 to 10 calves, and, bred even to the same bull, or to the same type bulls, it may not be till the 9th or 10th calf that she diverges from uniformity in her offspring. Thus one cow may throw 8 white calves and only the 9th be red or roan. This type of experience is not uncommon, and breeders not infrequently attribute it to the effect of age on the cow. To those who believe in individual prepotency—a view hardly consistent with Mendelism—this influence of age may not appear impossible. In the next place, the prejudice against white crosses on the part of breeders, indeed the dislike to white shorthorns in general, renders it almost impossible to effectively test the purity of any (RR) or (WW) bull, although his offspring may be fairly numerous. Further, the bulls are often mated solely within their particular owner's strain, and it is not so evident what would happen if they were mated outside it. This tendency to assortative mating among shorthorn breeders will be discussed more fully in another section. We have seen that there is historically a possibility of two strains of red and two strains of white having been mingled in the shorthorn. Determinants representing parti-colour and white markings can undoubtedly be introduced also—we confess to having made an attempt from this standpoint, which shattered with further examination of Table I.—but the introduction needs a wider practice than we can boast of in inventing Mendelian formulae, and until we are more convinced than we are at present of the soundness of such formulae we should prefer to leave the invention to those who have had it. *Coates' Herdbook* presents a wide range of material, and whatever we may think of the categories selected, the record has been made by persons in absolute ignorance of recent controversies about heredity. It is therefore really impartial material for Mendelians to unravel.

For the remainder of this paper we shall confine ourselves to the biometric treatment of the problem, which, as we have already noted, is independent of the existence or non-existence of any Mendelian formula. We shall compare the results with other investigations of a like kind and endeavour to interpret the results in biometric language.

(7) *Assortative Mating in the Shorthorn Population.*

Three sets of matings were worked out, namely: (*a*) the very recent matings of the parents of calves entered on the pedigree schedules, (*b*) the matings of the

great-great-grandparents of the calves on these schedules, and (*c*) a special series of matings taken from the early volumes of the *Herdbook*. Thus we have measures of the extent to which breeders now, 15 or 20 years ago, and perhaps 50 years ago, have mated like with like. The Tables each contain about 825 matings, and are given as Tables A, B, C of the Appendix. The general measure of the deviation from random mating was obtained by calculating the mean square contingency coefficient. The results were as follows:

Table of Assortative Mating Coefficients.

Recent Times:	·2030
15—20 Years ago:	·1912
50 Years ago:	·1219

It will thus be seen that the tendency of breeders to mate like with like has been very sensible for many years, and has now reached a value comparable with the unconscious assortative mating in the case of man*, or with that which breeders of fancy mice consciously or unconsciously produce†. This selective action ought to be taken into account when we endeavour to estimate the resemblance of parent to offspring.

(8) *On Direct Inheritance. First Degree.*

The usual four tables, sires and dams with bull and cow calves, were formed. These tables will be found as D, E, F, and G of the Appendix. They were worked out by mean square contingency, and, for the sake of comparison and control, also by mean contingency and by fourfold tables. In order to provide material, on which it might be possible to consider Mendelian formulae, Tables II. and III. are provided, giving both sire and dam for over 1000 bull and 1000 cow calves respectively. These tables have been already used in the previous section. They differ from the mating Table A, in that the same cow and bull are repeated not more than three times, while in the mating tables they are entered as often as they occur on the schedules.

A word must be said here about the contingency method as applied to these tables. The coefficient of mean square contingency is an absolute measure of the dependence of the offspring upon the parents, as long as we stick to the five groups now dealt with, but it would most probably be increased in value if we subdivided these groups—for example, divided the reds and roans into light and dark reds and roans. It has been pointed out elsewhere that this subdivision, without we immensely increase the number of individuals dealt with, must be limited, or we shall obtain very large contributions to the contingency from isolated units, which cannot of course be subdivided. Experience in comparing tables reduced by different methods seems to indicate that 25- to 49-fold contingency tables approach closely in average value to the results reached by 4-fold table methods‡. It would have been a great advantage in the case

* *Biometrika*, Vol. III. pp. 373 and 487. † *Biometrika*, Vol. III. p. 20.

‡ *Biometrika*, Vol. III. p. 253.

of the shorthorn tables to have been able to compare the 25-fold table results with those for 36- and 49-fold tables; and thus determine how far we were approximating to a limit. This would be the more important if we had any method of differentiating the large classes of reds and roans, which we have seen may quite possibly have sprung from two original races. As no sub-classification is provided by the *Herdbook*, this cannot be attempted, and we are left with the

TABLE II.

Distribution to Sires and Dams of ♂ Offspring.

Coat-Colour of Sire.

Coat-Colour of Dam.	Offspring	Red	Red, Little White	Red and White	Roan	White	Totals
Red	R. ...	75	10	5	68	1	159
	R. L. W....	10	4	3	13	—	30
	R. W. ...	5	3	—	4	1	13
	Ro. ...	4	1	1	91	27	124
	W. ...	—	—	—	2	—	2
	Group Total	94	18	9	178	29	328
Red, Little White	R. ...	6	2	—	8	—	16
	R. L. W....	9	1	1	5	—	16
	R. W. ...	3	1	2	3	—	9
	Ro. ...	—	1	—	15	6	22
	W. ...	—	—	—	—	—	—
	Group Total	18	5	3	31	6	63
Red and White	R. ...	14	2	3	10	1	30
	R. L. W....	8	1	1	8	—	18
	R. W. ...	5	2	9	10	—	26
	Ro. ...	3	2	—	40	9	54
	W. ...	—	—	—	—	—	—
	Group Total	30	7	13	68	10	128
Roan	R. ...	56	7	6	45	—	114
	R. L. W....	12	5	—	18	—	35
	R. W. ...	14	3	—	13	—	30
	Ro. ...	79	25	14	142	7	267
	W. ...	—	—	—	53	5	58
	Group Total	161	40	20	271	12	504
White	R. ...	—	—	—	1	—	1
	R. L. W....	1	—	—	1	—	2
	R. W. ...	—	—	—	—	—	—
	Ro. ...	26	4	3	18	—	51
	W. ...	—	—	—	8	—	8
	Group Total	27	4	3	28	—	62
	Grand Total	330	74	48	576	57	1085

TABLE III.

Distribution to Sires and Dams of ♀ Offspring.

Coat-Colour of Dam.	Offspring	Coat-Colour of Sire. Red	Red, Little White	Red and White	Roan	White	Totals
Red	R. ...	81	13	5	58	—	157
	R. L. W....	13	8	3	7	—	31
	R. W. ...	6	—	3	11	—	20
	Ro. ...	2	—	4	76	13	95
	W. ...	—	—	—	2	—	2
	Group Total	102	21	15	154	13	305
Red, Little White	R. ...	12	6	—	14	—	32
	R. L. W....	4	5	—	5	—	14
	R. W. ...	2	3	—	5	—	10
	Ro. ...	1	—	1	28	4	34
	W. ...	—	—	—	—	—	—
	Group Total	19	14	1	52	4	90
Red and White	R. ...	17	2	3	10	—	32
	R. L. W....	4	3	—	7	—	14
	R. W. ...	18	8	9	11	—	46
	Ro. ...	1	3	1	41	9	55
	W. ...	—	—	—	—	—	—
	Group Total	40	16	13	69	9	147
Roan	R. ...	61	12	8	41	—	122
	R. L. W....	14	4	3	13	—	34
	R. W. ...	10	6	5	22	—	43
	Ro. ...	78	17	17	136	7	255
	W. ...	—	—	—	31	4	35
	Group Total	163	39	33	243	11	489
White	R. ...	—	—	—	—	—	—
	R. L. W....	—	—	—	—	—	—
	R. W. ...	2	—	1	1	—	4
	Ro. ...	19	3	5	15	—	42
	W. ...	—	—	—	7	3	10
	Group Total	21	3	6	23	3	56
	Grand Total	345	93	68	541	40	1087

somewhat overweighted groups of red and roan—amounting to nearly $\frac{1}{3}$ and $\frac{1}{2}$ of the total offspring—to deal with. These large sub-groups are likely to reduce the mean square contingency coefficients, and we have no means of really testing how far we have already approached a limiting value of the mean square contingency coefficient.

In dealing with the material by the 4-fold table method, we considered that the best division we could make was between beasts with some whole red and

beasts with no whole red, the latter class including roans and whites. Table IV. gives our results:

TABLE IV.

Parental Inheritance in Shorthorns.

Nature of Percentage	Mean Square Contingency	Mean Contingency	Fourfold Table
Sire and Bull Calf ...	·37	·43	·46
Sire and Cow Calf ...	·40	·50	·52
Dam and Bull Calf ...	·40	·44	·44
Dam and Cow Calf ...	·42	·46	·43
Mean	·40	·46	·46

The result seems to us to indicate that the values by mean square contingency have come out somewhat too small. The general effect of an assortative mating if there were no selection of parents would be to increase the apparent parental correlation. If r were the real correlation, ρ the assortative mating, and $\mathfrak{r}$ the observed parental correlation, we should have

$$\mathfrak{r} = r\,\frac{1+\rho}{\sqrt{1+2\rho r^2}}.$$

If $r = \cdot5$ and $\rho = \cdot2$, we should expect the correlation altered by about $\frac{1}{10}$ increment. But this assumes that there has been no selection of parents, whereas in the case of the bulls at least there is obviously very considerable selection. The tables show that the red whites of both kinds are rejected, and that red and roan bulls are the most fashionable. But we know that any stringency in parental selection is more effective in reducing parental correlation than assortative mating can be in raising it*. The amount of this cannot be determined in the present case, but it very probably contributes to the reduction observed in the above table below the values determined for horse and greyhound.

The general conclusion that we can draw is that the parental resemblance to offspring in shorthorn cattle exceeds ·4, and may quite reasonably equal ·46, or even a higher value. It lies therefore well within the range of values ·4 to ·5 found already for man, horse, and dog.

(9) *On Direct Inheritance. Second and Higher Degrees.*

We now turn to the degree of resemblance between the offspring and more distant ancestry. The following results were obtained†:

* "Mathematical Contributions to the Theory of Evolution," *Phil. Trans.* Vol. 200 A, p. 39.

† The tabulated material will be found in the Appendix as Tables H, I, J, K, L, M, N, and O. No individual grandparent was taken more than six times, and on the average no individual occurs more than three times.

TABLE V.

Grandparental Inheritance in Shorthorns.

		Sex change
Paternal Grandsire and Bull Calf ...	·19	0
Paternal Grandsire and Cow Calf ...	·25	1
Maternal Grandsire and Bull Calf ...	·18	2
Maternal Grandsire and Cow Calf ...	·17	1
Paternal Grandam and Bull Calf ...	·15	1
Paternal Grandam and Cow Calf ...	·16	2
Maternal Grandam and Bull Calf ...	·24	1
Maternal Grandam and Cow Calf ...	·25	0
Mean =	·20	

It will be seen at once that this is lower than the results obtained for man, horse, and dog; i.e. about ·2 instead of ·3. This is what we might expect, as the parental correlations found for the shorthorn by mean square contingency are about ·4 instead of the ·5 for pigmentation in man, horse, and dog. We are not prepared to assert, however, that it connotes an actual lessening of the intensity of inheritance, because other methods in the parental case seem to indicate that the resemblance is really greater than is shown by mean square contingency with only a classification into five very unequal groups.

The values are clearly somewhat irregular. If we group them according to the number of sex changes we find:

Grandparental resemblance with no sex change = ·22
" " one " = ·20
" " two sex changes = ·17

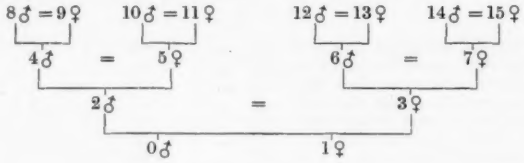
It would thus appear that, as in the case of eye-colour in man, there is a weakening of the intensity of resemblance with each change of sex. But if this be more than an *appearance*, it must be screened by other factors in the case of parentage, for in that case the values are ·40 for no change of sex and ·40 for one change of sex. This result is remarkably consonant with what we have previously observed in the case of greyhounds* where we found no sensible influence of change of sex in the parental resemblance, but an apparent influence on the grandparental relation. In eye-colour in the case of man it was found throughout all relationships†. We failed to demonstrate a change of sex influence in the case of coat-colour in horses‡, and it seems worth considering the matter further in shorthorns. Accordingly in attempting to estimate the great-grandparental influence, when we gave up the excessive labour of forming the 16 great-grandparental tables which would have meant dealing with upwards of 16,000 entries, and a far

* *Biometrika*, Vol. III. p. 258.

† Pearson: *Phil. Trans.* Vol. 195 A, pp. 114–117. Lutz: *Biometrika*, Vol. II. pp. 237–240.

‡ *Ibid.* Vol. II. p. 232.

larger number than this of pedigree great-grandparental relations*, we have simply classified into great-grandparents and (i) male, (ii) female descendants, with 0, 1, 2, and 3 changes of sex. This required the preparation of 8 instead of the full 16 tables. If we write the pedigree in the following manner:



then:

8 with 0	gave Table	P	of	great-grandsire and ♂	offspring,	no sex change
15 with 1	„	Q	„	grandam and ♀	„	„ „
8, 12, 14 with 1	„	R	„	grandsire and ♀	„	one „
9, 11, 15 with 0	„	S	„	grandam and ♂	„	„ „
10, 12, 14 with 0	„	T	„	grandsire and ♂	„	two changes
9, 11, 13 with 1	„	U	„	grandam and ♀	„	„ „
10 with 1	„	V	„	grandsire and ♀	„	three „
13 with 0	„	W	„	grandam and ♂	„	„ „

The results obtained, still by the method of mean square contingency, are given in the table below:

TABLE VI.

Great-Grandparental Inheritance in Shorthorns.

	Sex change	Resemblance
Great-Grandsires and ♂ Offspring ...	0	·20
„ „ „ „ ...	2	·12
„ „ ♀ „ ...	1	·18
„ „ „ „ ...	3	·17
Great-Grandams and ♂ Offspring ...	1	·15
„ „ „ „ ...	3	·15
„ „ ♀ „ ...	0	·26
„ „ „ „ ...	2	·12
	Mean = ·17	

These results are irregular, and not very conclusive; probably large enough numbers were not taken in several of the cases, but the tabulation of great-grandparents in large numbers is by no means an easy task, as in a pedigree stock high ancestors tend to be repeated many times. But in both grandsires and grandams, if we do not club together ♂ and ♀ offspring we find that no change of sex is

* The process throughout has been the same in all the ancestral tables; the produce is taken once only, and the same ancestor occurs on an average not more than three times, although individuals occur much more frequently, because as we go back there is much repeated blood owing to the fashion in bulls, and it would be almost impossible labour to reach 1000 independent great-great-grandsires. This method was adopted to avoid overweighting the tables with any particular individuals. The same principle was also used in the horse tables.

more potent than two changes (and than either one or three). There does not, however, seem to be much marked difference between one and three changes. On the whole it seems probable that a change of sex does produce some slight lessening in the ancestral resemblance beyond the parents; it is marked in eye-colour in man, and just appears in the coat-colours of greyhound and shorthorn, although if it exists in coat-colour of horses it has been screened by the probable errors of the results.

The mean result for eye-colour in man for great-grandparental resemblance is ·19, and for horse's coat-colour ·20, both results sensibly higher again than that for shorthorns.

Still one further ancestral resemblance has been worked out, namely, that between great-great-grandparents and the offspring. But the complete set of 32 tables has been replaced by four dealing only with each sex in the ancestry and the offspring regardless of the exact line of descent. The results are given below*:

TABLE VII. *Great-Great-Grandparent and Offspring.*

Great-Great-Grandsire and ♂ Offspring ...	·09
Great-Great-Grandsire and ♀ Offspring ...	·14
Great-Great-Grandam and ♀ Offspring ...	·13
Great-Great-Grandam and ♂ Offspring ...	·13
	Mean = ·12

The only comparable result is the combined table for all sexes provided by Dr Alice Lee† in the case of coat-colour in horses, and this gave for the great-great-grandparent and offspring the value ·15, thus agreeing with the previous cases in giving a higher value than the shorthorn results do.

We can now put together our results for ancestry in shorthorns and place alongside them those for other pigmentation researches:

TABLE VIII. *Ancestral Resemblance.*

Species	Man‡	Horse‡	Basset§ Hound	Greyhound‖	Shorthorn	Mean Result
Character	Eye-Colour	Coat-Colour	Coat-Colour	Coat-Colour	Coat-Colour	Pigmentation
Method employed	Fourfold Table	Fourfold Table	Fourfold Table	36-fold Contingency	25-fold Contingency	Various
Parent	·49	·52	·53	.53	·40	·49
Grandparent	·32	·30	·22	·33	·20	·27
Great-Grandparent ...	·19	·19	—	—	·17	·18
Great-Great-Grandparent	—	·15	—	—	·12	·13

* Given as Tables X, Y, Z, Ω, in the Appendix. † *Biometrika*, Vol. II. p. 235. ‡ *Ibid.* Vol. II. p. 222.
§ *R. S. Proc.*, Vol. LXVI. p. 157, omitting the cases of inheritance through the sire, as there can be little doubt that the sires are unreliable. ‖ *Biometrika*, Vol III. pp. 254–8.

But we do not consider these results as they stand as *strictly* comparable. We believe the lower values in the case of the shorthorns are not due to a lesser intensity of inheritance, but to the fact that the only possible contingency table is not finely enough divided. This we hold is indicated by the fourfold and mean contingency methods giving sensibly higher values than the mean square contingency method (see p. 448).

It appeared desirable to determine the best geometrical series: a, ar, ar^2, ar^3, to fit the shorthorn results. Accordingly we made:

$$(\cdot40 - a)^2 + (\cdot20 - ar)^2 + (\cdot17 - ar^2)^2 + (\cdot12 - ar^3)^2$$

a minimum. This led to the least square type equations:

$$\cdot40 + \cdot20r + \cdot17r^2 + \cdot12r^3 = a\,(1 + r^2 + r^4 + r^6),$$
$$\cdot20 + \cdot34r + \cdot36r^2 \qquad\quad = a\,(r + 2r^3 + 3r^5).$$

Eliminating a, r is given by:

$$\cdot17r^7 + \cdot28r^6 + 1\cdot20r^5 - \cdot04r^4 + \cdot63r^3 - \cdot36r^2 + \cdot06r - \cdot20 = 0.$$

The appropriate root of this is $r = \cdot64$ nearly, giving $a = \cdot382$, and the best fitting geometrical series is ·38, ·24, ·16, ·10, as against the observed ·40, ·20, ·17, ·12. But the remarkable result of the investigation is that the factor of reduction for the geometrical series in the case of eye-colour in man was found* to be ·62, and for coat-colour in horses ·66†, and the reducing factor for shorthorns is ·64, exactly their mean. The reducing factor for greyhounds is seen to be about ·62‡. Thus disregarding absolute values we see that the manner in which ancestral resemblance decreases with ascent is very closely the same for such different species as man, horse, greyhound, and shorthorn, and for most purposes may be taken equal to the round number 2/3. If we assume $r = 2/3$, we find the best value of a for the shorthorn is almost exactly ·37, giving the series ·37, ·25, ·16, and ·11, against the observed ·40, ·20, ·17, and ·12. It would thus appear that the general effect of applying the 25-fold mean contingency method has not been to lower the relative intensity of the successive grades of ancestry, but to weaken in about the same proportions the absolute values. The factor of reduction is about 38/50, or we should have for shorthorns the best fitting series,

·50, ·32, ·21, ·13,

as against an observed series,

·53, ·26, ·22, ·16.

Either or both of these are in quite reasonable agreement with what has been observed for man and horse.

Generally we may conclude that ancestral influence in the shorthorn diminishes at the same rate sensibly as in man, horse, and dog, and that absolute values are

* *Biometrika*, Vol. II. p. 222.

† *Ibid.* Vol. II. p. 222.

‡ The mean of Dr Warren's rather irregular results for *Hyalopterus* and *Daphnia* is again ·61. *Biometrika*, Vol. I. p. 139.

probably closely alike in the four cases, although the categories in the case of the shorthorn render the comparison harder.

(10) *On Collateral Inheritance.*

The work of tabling pairs of whole siblings is somewhat lengthy in the case of shorthorns, the breeding life of a cow is short, and there is a good deal of changing in the case of the bull used. Accordingly we have contented ourselves with the cases of whole brothers and whole sisters*. It is much easier to table cases of half siblings, and six tables of half siblings—brother-brother, sister-sister, brother-sister, the bull parent being common, and again the cow parent being common—have been constructed†. These half-sibling tables are of peculiar interest, because there is a good deal of obscurity at present about the degrees of likeness we should expect to find between half siblings. *A priori* we should expect the degree of likeness of half siblings to be half that of whole siblings, but this relationship has not been verified on any of the material hitherto considered: half siblings are relatively more alike than whole siblings, the addition of another common parent does not double the resemblance.

In Table IX below we have put together the results with regard to whole and half siblings as far as pigmentation results have yet been worked out.

Several conclusions may be drawn from this table. In the first place we see that half siblings through the bull are more alike than half siblings through the cow, and this is so for all three combinations. It would thus appear that the bull is more influential than the cow, but there is no prepotency of the bull over the cow in the parental resemblances. We should thus have the paradox that the bull caused the offspring to be more alike than the cow does, but not by causing them to be more like himself! Notwithstanding that the whole-sibling resemblances in shorthorns are of the same general value as those for eye-colour in man or coat-colour in Basset Hounds or Greyhounds of different litters, our opinion is that they are too low. Our experience with the parental tables leads us to believe that the grade of the resemblance is nearer that for hair-colour in man or coat-colour in horses. But we should not expect the reduction to be as great in this as in the parental case, owing to the special emphasis breeders lay on certain classes of parents. It is quite clear that neither in horses nor in shorthorns is the degree of resemblance for half siblings half that between whole siblings. The ratio of resemblance of maternal half siblings (the only case worked out for horses) to whole siblings is for horses ·57, and for shorthorns is ·56. These results are closely in accord, and our impression that a constant ratio exists between half and whole sibling resemblance, and that this ratio is greater than ·5, is not shaken by the values in the case of Basset Hounds. We think that a very profitable line of investigation would deal with the relation of half to whole brethren in a variety of characters, and, if possible, with measurable characters as well as pigments. So far as we can see in this, as in other features of collateral inheritance, shorthorns range themselves well within the limits already found for other species.

* Tables α and β of the Appendix. † Tables γ, δ, ϵ, ζ, η, θ, of the Appendix.

TABLE IX.

Collateral Inheritance.

Species	Man			Horse	Shorthorn	Basset Hound	Greyhound		Mean
Character	Eye-Colour		Hair-Colour	Coat-Colour	Coat-Colour	Coat-Colour	Coat-Colour, Studbook Returns		Pigmentation
	I	II					36-fold Contingency		
Reduction Method	4-fold Table	4-fold Table	4-fold Table	4-fold Table	25-fold Contingency	4-fold Table	Same Litter	Different Litter	Various
Whole Brothers	·52	·54	·62	·62	·52	·53	·52	·52	·55
Whole Sisters ...	·45	·52	·57	·69	·54		·60	·56	·56
Whole Brother and Sisters	·46	·53	·55	·58	—		·56	·51	·53
Half Brothers through Father	—	—	—	—	·34	·22* (All half siblings through Dam)	—	—	·34
Half Brothers through Mother	—	—	—	·36	·30		—	—	·33
Half Sisters through Father	—	—	—	—	·36		—	—	·36
Half Sisters through Mother	—	—	—	·43	·29		—	—	·36
Half Sister and Brother through Father	—	—	—	—	·32		—	—	·32
Half Sister and Brother through Mother	—	—	—	·28	·30		—	—	·29
Mean result for whole siblings	·48	·53	·58	·63	·53	·53	·56	·53	·55
Mean result for half siblings	—	—	—	·36	·32	·22*	—	—	·33

(11) *General Conclusions.*

(*a*) With the exception of blue-grey cattle we are hardly likely to meet with a simpler colour combination in cattle than that to be found in the case of the shorthorn. If we consider the case of the cross of the polled black Galloway and the white shorthorn, then whether we deal with coat-colour or horns we find certain phases which suggest Mendelian ratios, but no simple Mendelian formula applies *rigidly*; we find ourselves neglecting sensible percentages of occurrences incompatible with a theory of the pure gamete. The hybrids of two races which at any rate breed approximately true, like the white shorthorn and white wild cattle, throw a considerable range of colours when they are mated. Thus to describe the colour action of shorthorn crosses will need a differentiation of the primary colour categories into homozygous and heterozygous groups, or the introduction of a

* Not included in the general means of half siblings, as of doubtful validity.

somewhat complex system of latent characters, because practically every colour group will be found—with larger or smaller frequencies—to reproduce every other colour group. The only alternative is to reject the records in bulk as patently in error, because they do not fit rigidly a simple Mendelian scheme.

(*b*) When we turn to the statistical treatment of the data we are at once met by a further difficulty. The method of contingency enables us to free ourselves from any question of colour order or scale, and thus we are not troubled with any hypothesis as to a normal or Gaussian distribution. But while thus taking its place as the rational statistical process for dealing with colour data, we are dependent on breeders' colour classifications which are not made with a view to statistical treatment, and provide the statistician, as in this case of shorthorns, with groupings like those of red and roan not fine enough for his purposes. He obtains measures of intensity of resemblance comparable within the species—he can investigate the relative intensity of ancestral and collateral likenesses, but it is not so easy to compare the results for one species with a second. It would be a great advantage if red and roan could be recorded in light and dark shades, for the range in these colours is very wide and needs differentiation. With a 36- or 49-fold table we should undoubtedly get results better comparable with those for horses and greyhounds.

(*c*) Notwithstanding these difficulties we actually find, however, that the shorthorns fit quite reasonably into the general range of results for man, horse, and dog. The "decay" of ancestral resemblance was found to be ·64, a value lying between that found for man and dog, and expressible in round numbers by saying that the relationship decreases 2/3 at each upward grade. Thus the geometrical series feature of the ancestral law is maintained in the case of shorthorns. The parental relationship lies between ·4 and ·5, and is probably nearer the latter than the former number; the whole sibling resemblance is a little over ·5 at a minimum, and appears to be very similar in range of value for that found for man and dog. There is a substantial assortative mating produced by breeders mating like with like, but while this would raise the apparent values of the contingency coefficients, the fact that red and roan are especially fashionable tends by its selective action to produce more than an equivalent lowering. Generally the shorthorn data strengthen our view that—whether it be Mendelian, or some other yet to be found—there is a single basal law of inheritance, manifesting itself notwithstanding assortative mating, selection, and environment in the clustering of the inheritance constants for all species of populations about the value ·45 to ·5 for parental inheritance, and about ·5 to ·55 for fraternal inheritance, with a geometrical series of a ratio roughly about 2/3 for ancestral grades. Such an ancestral law will, we are convinced, be found to flow from the physiological basal law if it be discovered. Meanwhile each fresh series as it is worked out biometrically tends to strengthen the view that apparent variations in the intensity of hereditary resemblance are not real variations in nature. The hereditary process is one, producing in a population definite degrees of average resemblance

between grades of kinship, and the apparent variation in intensity of inheritance is due to environmental influence, the selective action of breeders, assortative mating, and the difficulty of applying a uniform method of reduction to categories so varying in their extent and frequency as those we meet with in dealing with the records of quantitatively undetermined characters like pigmentation.

APPENDIX.

TABLES OF SHORTHORN MATING AND INHERITANCE.

I. ASSORTATIVE MATING PRODUCED BY BREEDERS' SELECTIONS.

A. *Modern Matings.*

Bull.

Cow.	R.	R. L. W.	R. W.	Ro.	W.	Totals
R. ...	69	18	7	129	23	246
R. L. W. ...	17	13	5	36	5	76
R. W. ...	12	7	4	43	10	86
Ro. ...	100	29	15	216	9	369
W. ...	16	4	5	20	3	48
Totals	224	71	36	444	50	825

B. *Matings of 15 to 20 Years ago.*

Bull.

Cow.	R.	R. L. W.	R. W.	Ro.	W.	Totals
R. ...	63	10	9	105	3	190
R. L. W. ...	11	4	1	19	5	40
R. W. ...	39	8	13	62	4	126
Ro. ...	142	30	40	188	18	418
W. ...	29	3	7	21	—	60
Totals	284	55	70	395	30	834

C. *Matings of 50 Years ago.*

Bull.

Cow.	R.	R. L. W.*	R. W.	Ro.	W.	Totals
R. ...	10	—	20	54	11	95
R. L. W. ...	—	—	—	—	—	—
R. W. ...	17	—	47	137	38	239
Ro. ...	38	—	64	292	41	435
W. ...	3	—	15	34	7	59
Totals	68	—	146	517	97	828

* Formerly recorded under R. W.

II. PARENTS AND OFFSPRING.

D. *Sires and Bull Calves.*

Sires.

Bull Calves.	R.	R. L. W.	R. W.	Ro.	W.	Totals
R. ...	151	21	14	132	2	320
R. L. W. ...	40	11	5	45	—	101
R. W. ...	27	9	11	30	1	78
Ro. ...	112	33	18	306	49	518
W. ...	—	—	—	63	5	68
Totals	330	74	48	576	57	1085

E. *Sires and Cow Calves.*

Sires.

Cow Calves.	R.	R. L. W.	R. W.	Ro.	W.	Totals
R. ...	171	33	16	123	—	343
R. L. W. ...	35	20	6	32	—	93
R. W. ...	38	17	18	50	—	123
Ro. ...	101	23	28	296	33	481
W. ...	—	—	—	40	7	47
Totals	345	93	68	541	40	1087

F. *Dams and Bull Calves.*

Dams.

Bull Calves.	R.	R. L. W.	R. W.	Ro.	W.	Totals
R. ...	159	16	30	114	1	320
R. L. W. ...	30	16	18	35	2	101
R. W. ...	13	9	26	30	—	78
Ro. ...	124	22	54	267	51	518
W. ...	2	—	—	58	8	68
Totals	328	63	128	504	62	1085

G. *Dams and Cow Calves.*

Dams.

Cow Calves.	R.	R. L. W.	R. W.	Ro.	W.	Totals
R. ...	157	32	32	122	—	343
R. L. W. ...	31	14	14	34	—	93
R. W. ...	20	10	46	43	4	123
Ro. ...	95	34	55	255	42	481
W. ...	2	—	—	35	10	47
Totals	305	90	147	489	56	1087

III. GRANDPARENTS AND OFFSPRING.

H. *Paternal Grandsire and Male Offspring.*

Paternal Grandsire.

♂ Offspring.	R.	R. L. W.	R. W.	Ro.	W.	Totals
R. ...	124	21	9	129	13	296
R. L. W. ...	32	7	4	50	3	96
R. W. ...	22	4	6	50	2	84
Ro. ...	143	19	24	276	32	494
W. ...	9	2	1	36	1	49
Totals	330	53	44	541	51	1019

I. *Paternal Grandsire and Female Offspring.*

Paternal Grandsire.

♀ Offspring.	R.	R. L. W.	R. W.	Ro.	W.	Totals
R. ...	160	19	14	115	7	315
R. L. W. ...	28	8	4	54	1	95
R. W. ...	39	7	11	42	2	101
Ro. ...	124	30	26	261	15	456
W. ...	14	1	4	34	4	57
Totals	365	65	59	506	29	1024

J. *Maternal Grandsire and Male Offspring.*

Maternal Grandsire.

♂ Offspring.	R.	R. L. W.	R. W.	Ro.	W.	Totals
R. ...	151	31	20	135	5	342
R. L. W. ...	39	6	10	49	2	106
R. W. ...	21	9	6	48	2	86
Ro. ...	180	34	38	285	26	563
W. ...	19	1	2	34	2	58
Totals	410	81	76	551	37	1155

K. *Maternal Grandsire and Female Offspring.*

Maternal Grandsire.

♀ Offspring.	R.	R. L. W.	R. W.	Ro.	W.	Totals
R. ...	126	28	23	161	13	351
R. L. W. ...	39	8	7	37	4	95
R. W. ...	29	10	15	40	1	95
Ro. ...	117	39	43	245	20	464
W. ...	9	5	5	39	3	61
Totals	320	90	93	522	41	1066

L. *Paternal Grandam and Male Offspring.*

Paternal Grandam.

♂ Offspring.	R.	R. L. W.	R. W.	Ro.	W.	Totals
R. ...	116	19	49	186	18	388
R. L. W. ...	27	5	15	50	4	101
R. W. ...	21	9	26	56	5	117
Ro. ...	119	20	61	277	44	521
W. ...	10	3	4	33	6	56
Totals	293	56	155	602	77	1183

M. *Paternal Grandam and Female Offspring.*

Paternal Grandam.

♀ Offspring.	R.	R. L. W.	R. W.	Ro.	W.	Totals
R. ...	123	13	29	176	18	359
R. L. W. ...	36	3	16	47	3	105
R. W. ...	22	4	12	51	5	94
Ro. ...	145	17	53	306	40	561
W. ...	14	—	4	28	7	53
Totals	340	37	114	608	73	1172

N. *Maternal Grandam and Male Offspring.*

Maternal Grandam.

♂ Offspring.	R.	R. L. W.	R. W.	Ro.	W.	Totals
R. ...	131	32	43	125	11	342
R. L. W. ...	29	6	18	34	3	90
R. W. ...	16	7	11	42	2	78
Ro. ...	127	33	72	287	32	551
W. ...	7	2	2	46	6	63
Totals	310	80	146	534	54	1124

O. *Maternal Grandam and Female Offspring.*

Maternal Grandam.

♀ Offspring.	R.	R. L. W.	R. W.	Ro.	W.	Totals
R. ...	133	22	46	130	10	341
R. L. W. ...	28	8	8	33	3	80
R. W. ...	25	6	26	43	8	108
Ro. ...	94	32	67	235	32	460
W. ...	6	4	—	32	5	47
Totals	286	72	147	473	58	1036

IV. GREAT-GRANDPARENTS AND OFFSPRING.

P. *Great-Grandsire and Male Offspring. No sex change.*

Great-Grandsire.

♂ Offspring.	R.	R. L. W.	R. W.	Ro.	W.	Totals
R. ...	82	4	6	89	5	186
R. L. W....	25	3	2	28	1	59
R. W. ...	14	0	4	30	2	50
Ro. ...	115	6	8	154	8	291
W. ...	6	4	1	19	1	31
Totals	242	17	21	320	17	617

Q. *Great-Grandam and Female Offspring. No sex change.*

Great-Grandam.

♀ Offspring.	R.	R. L. W.	R. W.	Ro.	W.	Totals
R. ...	52	10	23	65	13	163
R. L. W....	8	7	5	17	1	38
R. W. ...	7	—	13	21	1	42
Ro. ...	48	11	35	103	18	215
W. ...	1	1	4	15	1	22
Totals	116	29	80	221	34	480

R. *Great-Grandsire and Female Offspring. One sex change.*

Great-Grandsire.

♀ Offspring.	R.	R. L. W.	R. W.	Ro.	W.	Totals
R. ...	220	30	32	207	10	499
R. L. W....	57	11	9	70	5	152
R. W. ...	47	7	10	82	4	150
Ro. ...	228	50	43	468	25	814
W. ...	21	6	10	60	6	103
Totals	573	104	104	887	50	1718

S. *Great-Grandam and Male Offspring. One sex change.*

Great-Grandam.

♂ Offspring.	R.	R. L. W.	R. W.	Ro.	W.	Totals
R. ...	187	30	84	262	22	585
R. L. W....	54	1	21	81	17	174
R. W. ...	36	1	14	72	12	135
Ro. ...	239	46	105	477	71	938
W. ...	24	4	10	59	5	102
Totals	540	82	234	951	127	1934

T. *Great-Grandsire and Male Offspring. Two sex changes.*

Great-Grandsire.

♂ Offspring.	R.	R. L. W.	R. W.	Ro.	W.	Totals
R. ...	263	36	33	230	15	577
R. L. W....	60	11	11	71	5	158
R. W. ...	47	10	8	78	4	147
Ro. ...	287	41	56	400	29	813
W. ...	30	8	5	42	4	89
Totals	687	106	113	821	57	1784

U. *Great-Grandam and Female Offspring. Two sex changes.*

Great-Grandam.

♀ Offspring.	R.	R. L. W.	R. W.	Ro.	W.	Totals
R. ...	172	27	73	282	32	586
R. L. W....	48	7	18	81	8	162
R. W. ...	42	9	38	79	7	175
Ro. ...	199	34	111	386	66	796
W. ...	25	3	21	52	3	104
Totals	486	80	261	880	116	1823

V. *Great-Grandsire and Female Offspring. Three sex changes.*

Great-Grandsire.

♀ Offspring.	R.	R. L. W.	R. W.	Ro.	W.	Totals
R. ...	69	17	17	70	2	175
R. L. W....	17	1	3	21	2	44
R. W. ...	18	2	5	21	3	49
Ro. ...	101	13	17	112	14	257
W. ...	16	—	3	17	1	37
Totals	221	33	45	241	22	562

W. *Great-Grandam and Male Offspring. Three sex changes.*

Great-Grandam.

♂ Offspring.	R.	R. L. W.	R. W.	Ro.	W.	Totals
R. ...	65	7	31	100	17	220
R. L. W....	12	5	11	29	3	60
R. W. ...	10	4	14	23	3	54
Ro. ...	87	12	39	163	22	323
W. ...	6	2	4	20	3	35
Totals	180	30	99	335	48	692

V. GREAT-GREAT-GRANDPARENTS AND OFFSPRING.

X. *Great-Great-Grandsire and Male Offspring.*

Great-Great-Grandsire.

♂ Offspring.	R.	R. L. W.	R. W.	Ro.	W.	Totals
R. ...	266	26	47	248	22	609
R. L. W....	64	3	9	55	5	136
R. W. ...	66	11	21	82	8	188
Ro. ...	315	32	55	374	30	806
W. ...	39	4	8	41	3	95
Totals	750	76	140	800	68	1834

Y. *Great-Great-Grandsire and Female Offspring.*

Great-Great-Grandsire.

♀ Offspring.	R.	R. L. W.	R. W.	Ro.	W.	Totals
R. ...	215	20	29	177	27	468
R. L. W....	43	3	11	61	7	125
R. W. ...	35	4	8	39	5	91
Ro. ...	250	16	63	318	39	686
W. ...	26	4	8	59	6	103
Totals	569	47	119	654	84	1473

Z. *Great-Great-Grandam and Male Offspring.*

Great-Great-Grandam.

♂ Offspring	R.	R. L. W.	R. W.	Ro.	W.	Totals
R. ...	236	25	105	364	40	770
R. L. W....	64	6	7	79	13	169
R. W. ...	46	5	36	116	16	219
Ro. ...	256	33	133	505	57	984
W. ...	29	9	23	50	8	119
Totals	631	78	304	1114	134	2261

Ω. *Great-Great-Grandam and Female Offspring.*

Great-Great-Grandam.

♀ Offspring.	R.	R. L. W.	R. W.	Ro.	W.	Totals
R. ...	169	19	68	258	27	541
R. L. W....	36	3	19	69	4	131
R. W. ...	19	5	17	58	15	114
Ro. ...	214	25	98	437	54	828
W. ...	24	3	18	53	11	109
Totals	462	55	220	875	111	1723

VI. COLLATERAL INHERITANCE. WHOLE AND HALF-SIBLINGS.

α. *Whole Brothers.*

First Brother.

Second Brother.	R.	R. L. W.	R. W.	Ro.	W.	Totals
R. ...	290	29	69	155	27	570
R. L. W. ...	29	44	8	22	1	104
R. W. ...	69	8	108	109	12	306
Ro. ...	155	22	109	494	75	855
W. ...	27	1	12	75	50	165
Totals	570	104	306	855	165	2000

β. *Whole Sisters.*

First Sister.

Second Sister.	R.	R. L. W.	R. W.	Ro.	W.	Totals
R. ...	302	26	48	155	18	549
R. L. W. ...	26	40	9	23	2	100
R. W. ...	48	9	112	98	17	284
Ro. ...	155	23	98	554	81	911
W. ...	18	2	17	81	38	156
Totals	549	100	284	911	156	2000

γ. *Half Brothers through Sire.*

First Half Brother.

Second Half Brother.	R.	R. L. W.	R. W.	Ro.	W.	Totals
R. ...	348	59	42	245	25	719
R. L. W. ...	59	54	19	49	10	191
R. W. ...	42	19	26	59	5	151
Ro. ...	245	49	59	446	58	857
W. ...	25	10	5	58	24	122
Totals	719	191	151	857	122	2040

δ. *Half Sisters through Sire.*

First Half Sister.

Second Half Sister.	R.	R. L. W.	R. W.	Ro.	W.	Totals
R. ...	266	41	41	246	17	611
R. L. W. ...	41	42	19	37	9	148
R. W. ...	41	19	40	65	20	185
Ro. ...	246	37	65	518	67	933
W. ...	17	9	20	67	24	137
Totals	611	148	185	933	137	2014

ε. Half Brother and Sister through Sire.

Half Brother.

Half Sister.	R.	R. L. W.	R. W.	Ro.	W.	Totals
R. ...	144	26	33	99	6	308
R. L. W. ...	17	22	4	24	6	73
R. W. ...	38	7	14	32	7	98
Ro. ...	136	25	33	215	41	450
W. ...	15	4	1	45	6	71
Totals	350	84	85	415	66	1000

ζ. Half Brothers through Dam.

First Half Brother.

Second Half Brother.	R.	R. L. W.	R. W.	Ro.	W.	Totals
R. ...	294	32	83	185	22	616
R. L. W. ...	32	12	18	41	14	117
R. W. ...	83	18	54	95	21	271
Ro. ...	185	41	95	428	80	829
W. ...	22	14	21	80	30	167
Totals	616	117	271	829	167	2000

η. Half Sisters through Dam.

First Half Sister.

Second Half Sister.	R.	R. L. W.	R. W.	Ro.	W.	Totals
R. ...	246	29	66	215	26	582
R. L. W. ...	29	22	7	42	10	110
R. W. ...	66	7	74	107	16	270
Ro. ...	215	42	107	416	91	871
W. ...	26	10	16	91	24	167
Totals	582	110	270	871	167	2000

θ. Half Brother and Sister through Dam.

Half Brother.

Half Sister.	R.	R. L. W.	R. W.	Ro.	W.	Totals
R. ...	136	17	31	114	12	310
R. L. W. ...	18	9	—	21	6	54
R. W. ...	26	9	31	38	12	116
Ro. ...	121	26	64	190	36	437
W. ...	17	2	5	41	18	83
Totals	318	63	131	404	84	1000

HEREDITARY DEAFNESS.

A DISCUSSION OF THE DATA COLLECTED BY DR E. A. FAY IN AMERICA.

By EDGAR SCHUSTER, M.A., Francis Galton Research Fellow in National Eugenics in the University of London.

THE present paper is an attempt* to apply the newer methods of statistics, introduced by Mr Francis Galton and developed by Prof. Karl Pearson, to a valuable collection of hereditary data that have hitherto only been discussed by the older ones.

The object is to obtain results which admit of direct comparison with others derived from quite other collections. These results take the form of "correlation coefficients," which give a numerical measure of the intensity of heredity.

The best material for this purpose would have been a random sample of the population, of which the following particulars were known, namely the facts which relate to the presence or absence of deafness, and of its nature when present, among their parents, children, brothers and sisters, etc.; and not only this but also the total number of their relatives in those degrees in which the number is variable.

Deafness of the degree with which we are dealing is fortunately of such rarity that a sample of this kind, in order to be useful, would have to contain at least a million persons, so that practically speaking it is outside the range of possible attainment for a private individual.

In default of this one is reduced to processes involving many possibilities of error, as one has to supply by estimate the deficiencies in such information as is accessible, in order to obtain artificial samples for making correlation tables.

It might be objected that the facts taken have already been stated in a form that anyone can understand, and are now presented in a way that is only intelli-

* I have to thank Prof. Pearson for giving me very much help in devising the special methods employed in this paper; my thanks are also due to Mr Galton for reading the manuscript and suggesting many improvements in its form and style.

gible to persons with a certain amount of special knowledge, and moreover, owing to the various errors which may have crept in during the process, gives uncertain and misleading information even to them. In answer one must lay stress on the extreme importance of bringing the facts relating to the inheritance of rare characters such as congenital deafness into line with those relating to common ones, such for instance as tallness or shortness of stature. Without such comparisons it is impossible to make real advance towards knowledge of the general laws of heredity, so that even an imperfect attempt at making them is better than none at all. The results need mislead no one, as every assumption and every known source of error is pointed out; and as statistical methods do not consume the material which they use, the same data can be employed over and over again to test any new theory which may seem worthy of investigation.

Material. The material on which this paper is based is derived entirely from E. A. Fay's "Marriages of the Deaf in America" (Volta Bureau, Washington, U.S.A., 1898, pp. vii and 527). The collection of the facts contained in this book was undertaken, as a labour of love, by Dr E. A. Fay, the editor of *The American Annals of the Deaf*, the expenses of the investigation being borne by the Volta Bureau, which had been endowed, partly for this purpose, by Dr A. Graham Bell. How much labour must have been needed in order to obtain this immense and valuable mass of data can be realised from the fact that nearly every marriage of a deaf person, which had taken place in America and Canada from 1850 to 1894, and a large number which took place before that period are here recorded. There are in all, included in the book, records of the marriages of 4471 pairs of persons, of whom at least one in each pair was known to be of a degree of deafness defined on p. 7 of the introduction in the following words:—

"The term 'deaf,' as used in this monograph, invariably refers to the class of persons for whose benefit our 'schools for the deaf' are intended—persons so deaf from birth, childhood or youth that they cannot be educated in common schools, persons who are more frequently but less accurately classed as 'the deaf and dumb' or 'deaf-mutes.' It includes all persons who are recorded in school reports, census reports, marriage records, etc., as 'deaf and dumb,' 'deaf-mutes,' 'totally deaf,' 'very deaf,' 'deaf,' or 'very hard of hearing,' and all who have attended the schools for the deaf. Persons recorded as 'slightly deaf,' 'partially deaf,' 'deaf in one ear,' 'deaf in adult life,' 'hard of hearing,' etc., are not included, unless they have attended schools for the deaf."

Pages 138–499 consist of a tabular record of the 4471 marriages, the details given concerning each being as follows:—Date of information, date of marriage, total number of children, number deaf, number hearing, number unknown whether deaf or hearing. After this come the following particulars concerning both the husband and the wife; whether they were deaf or hearing or whether this was unknown, the year of birth, the total number of the family in which they came (i.e. their brothers and sisters and themselves); how many of these were deaf, how many hearing, and about how many this was not reported; whether they had any

deaf relatives excluding brothers, sisters and children; whether the deafness was congenital or if not at what age it occurred, and to what cause it was attributed; the schools at which they were educated, and lastly there is a broad space left for remarks. The remarks consist mostly of indicating whether the husband or wife was related to, or had been married to, or was subsequently married to, any other of the husbands or wives, whose marriages are recorded in the tables. Where any of these particulars could not be obtained, a dotted line is ruled across the space in the table in which it should have been set down. The first part of the book consists of an analysis of the tables, but for various reasons I have not made any use of this part.

Deafness. Deafness may be divided into two categories, congenital and "acquired" or "adventitious." Congenital deafness appears for the most part to be due to faults in the structure of the membranous labyrinth of the inner ear; but the eighth nerve may also be affected, and in some cases the deafness results from defects in the central nervous system, which may coexist with perfectly formed ears. It is usually associated with an anomalous condition of the functions of balance, and much experimental work has been done recently in Germany on this point (by Pollak, Alexander, Kreidl, Hammerschlag, etc.). It may possibly be compared to the structural peculiarities of the semicircular canals of Japanese waltzing mice.

Acquired deafness is due generally to changes in the middle and inner ears produced by inflammation, which may be idiopathic or set up by various diseases, such as meningitis, scarlet fever, etc.

Should a child be born with the faculty of hearing not developed beyond a certain pitch, he will be unable to learn the art of speech in the ordinary manner, that is to say by imitating the sounds formed by the speech of others; and should he lose his hearing before, or but shortly after, learning to talk, the effect will be the same. He may, however, be taught to speak in a fairly intelligible manner by special methods, which consist in causing him to imitate not the sounds themselves, but the different muscular movements that produce them. But in order that this may be done with success, there needs to be patience and skill on the part of the teacher and fair mental ability on the part of the pupil.

In making the correlation tables contained in this paper, it has been impossible to separate congenital from acquired deafness, because one is only given information on this point for one set of people, as will be explained later, and not for their brothers or sisters or children. Also it is almost certain that a large number of congenital deaf-mutes are included in the adventitious class, as this is apparently considered a matter of less shame than to be born deaf. If a child is found to be deaf at the age when it should begin to speak, it is very difficult to say definitely whether it has been born deaf, or whether the deafness was produced by some disease in its infancy, so that when there is any possible pretext for attributing it to the latter cause, this is done. As the possession or non-possession of deaf

relatives is sometimes used to assist classification into the congenital and acquired classes, it follows that those who are included in the former class cannot be used as a fair sample of true congenital deaf-mutes for the purpose of the study of the inheritance of deafness.

Methods and Results.

(1) *Correlation between Father and Children and Mother and Children.*

The first step taken was to tabulate afresh the offspring born from the various classes of marriages enumerated below, in doing which I have only used such families as had their total number of offspring definitely stated. Table I. shows

TABLE I.

Showing number of marriages and number of deaf and hearing offspring produced from them, in which either the father or the mother or both parents were deaf.

	Class of Marriage	Number of Marriages	Number of Children		
			Deaf	Hearing	Doubtful
1	Father born deaf, Mother born deaf	311	184	480	46
2	Father born deaf, Mother not born deaf	552	93	1078	84
3	Mother born deaf, Father not born deaf	562	96	936	86
4	Father deaf but not known to be born deaf, Mother ditto ...	1142	68	1696	217
5	Father deaf but not known to be born deaf, Mother born deaf	478	72	767	82
6	Mother deaf but not known to be born deaf, Father born deaf	442	66	831	70
7	Father deaf but not known to be born deaf, Mother not deaf	234	47	449	21
8	Mother deaf but not known to be born deaf, Father not deaf	153	15	292	18
9	Total of marriages in which Father was deaf	2717	464	4470	450
10	Total of marriages in which Mother was deaf	2610	429	4235	437

the result of this process. The first class of marriage consists of those in which both the husband and wife (the father and mother of the family) were stated to be congenitally deaf. The second consists of those in which the husband was congenitally deaf but the wife was either adventitiously deaf, not deaf at all, or where doubt existed as to nature of her deafness, or as to whether she was deaf or not. Class 3 is the converse of this. In class 4 both the husband and wife were deaf, but they were either adventitiously deaf or at any rate were not certainly known to be congenitally deaf. Class 5: the husband's deafness was similar in nature to that of class 4, but the wife was born deaf, and class 6 is the converse of class 5. In classes 7 and 8 the husband in the former, and the wife in the latter, were deaf in the same manner as those in class 4, while the wife in the first case, and the husband in the last, were not deaf at all. Class 9 sums up all the cases in which the husband was deaf at all, and class 10 does the same for the wives.

It is interesting to note the high proportion (·2771 of the number about whom it was known whether they were deaf or hearing) of deaf children in class 1

as compared to class 4 (·0386), while classes 2 and 3 are intermediate between the two with ·0794 and ·0930 respectively. I do not however intend going into this point, as it is dealt with by Fay in some detail. I have not made use of Fay's tables principally because no distinction of the sexes was made, fathers and mothers were grouped together, and I thought it advisable to keep them separate to see whether the paternal correlation was different from the maternal.

The next step was to tabulate all those cases in which either the father or the mother of the husbands and wives of the marriages recorded was deaf. For the result of this process see Table II. The total number of husbands and wives in

TABLE II.

Showing the whole number of husbands and wives about whom definite information is given as to whether or no they had any deaf relatives other than children and brothers and sisters, also what number of these had either one or both parents deaf.

	Congenital				Adventitious			
	Total Number	Father Deaf	Mother Deaf	Both Parents Deaf	Total Number	Father Deaf	Mother Deaf	Both Parents Deaf
Husbands ...	648	11	7	37	1147	4	3	13
Wives ...	672	12	13	36	1184	5	3	14

this table means the total numbers about whom definite information was given as to whether or not they had deaf relatives other than children or brothers and sisters. The numbers in the columns "Father deaf," "Mother deaf," "Both parents deaf," possibly do not represent the whole numbers about whom this statement could have been made truthfully; but they must be looked upon as giving minimum values for them.

It was from Tables I. and II. that the paternal and maternal correlation tables were constructed. As it was only possible to divide the material into two classes, namely deaf and hearing, the ordinary form of correlation table could not be used, and the fourfold method had to be adopted. This method involves an assumption, which in the present instance it is impossible to test, that the distribution of the material with regard to the variable character dealt with is normal. Where this is the case, or even approximately the case, it has been shown that one obtains practically the same results from this form of table as one does from the longer one.

The actual process of constructing the correlation table was as follows: firstly, from Table I. one learns that 2717 deaf fathers (classing those that were born deaf and those that became deaf after birth together) have 464 deaf children and 4470 hearing children, also 450 children about whom it was not known

whether they were deaf or hearing. There is no reason to suppose that among the doubtful ones the deaf did not bear the same proportion to the hearing as among the others; that is to say that among 450 of them there would be 42 deaf and 408 hearing; therefore by adding these to the others we have 506 deaf and 4878 hearing. Weighting the fathers with the number of their children (i.e. including them once for each child that they have), there are then 4878 deaf fathers with hearing children, and 506 with deaf. In order to complete the table it is necessary to know how many hearing fathers with hearing children, and how many with deaf children, would occur in a random sample of the population containing the deaf fathers mentioned above. Now from Table II. we know that a total of 3658 deaf people have in 132 cases deaf fathers, and in the remaining 3519 cases hearing fathers; and we have to include in the table we are constructing 506 deaf people with deaf fathers, therefore we may take it that the number of deaf people with hearing fathers which correspond with this number is 506/132 multiplied by 3519, which is equal to 13,489.

There only remains in order to complete the table to determine how many hearing people with hearing fathers should be included in the sample. We know from the American census of 1890 that 659 deaf people occur in every million of the general population. We are dealing here with 506 plus 4878 = 5384 deaf fathers, therefore if we assume that deaf fathers bear the same numerical relation to hearing fathers, as deaf persons in general bear to hearing persons, then the 5384 deaf fathers occur among a total number of 5384/659 multiplied by a million, which is equal to 8,169,954 fathers, and as 5384 of these are deaf there remain 8,164,570 who are hearing. Of these 13,489 have deaf offspring, therefore 8,151,081 have hearing offspring, and thus Table III. may be completed.

TABLE III.

Fathers.

Children.	Deaf	Hearing
Hearing ...	4878	8,151,959
Deaf... ...	506	13,489

$h = 2{\cdot}926730$
$H = {\cdot}005507$
$k = 3{\cdot}212070$
$K = {\cdot}002294$
$r = {\cdot}51$

TABLE IV.

Fathers.

Children.	Deaf	Hearing
Hearing ...	4878	21,222,727
Deaf... ...	506	13,489

$h = 3{\cdot}477260$
$H = {\cdot}000945$
$k = 3{\cdot}212070$
$K = {\cdot}002294$
$r = {\cdot}57$

There is also another way of completing the table. If instead of assuming that deaf fathers bear the same numerical relation to hearing fathers, as deaf persons in

general do to hearing persons, we work on the supposition that this relation is borne by deaf children to hearing children; then as we are dealing with 13,489 plus 506 = 13,995 deaf children, these must occur among a total number of 21,236,722 children, of whom 21,222,727 are hearing, and as 4878 have deaf fathers, 21,217,849 would have normal fathers, and in this way Table IV. may be constructed.

TABLE V.

Fathers.

Children.	Deaf	Hearing
Hearing ...	4878	8,159,692
Deaf... ...	506	4878

$h = k = 3{\cdot}212070$
$H = K = {\cdot}002294$
$r = {\cdot}62$

TABLE VI.

Fathers.

Children.	Deaf	Hearing
Hearing ...	3519	5,533,042
Deaf... ...	132	3519

$h = k = 3{\cdot}212070$
$H = K = {\cdot}002294$
$r = {\cdot}46$

Tables III. and IV. thus combine the information derived from Tables I. and II. In Table II. the assumption is made that the proportion of deaf to hearing fathers is that of the general population, and in Table IV. a similar assumption is made for the children. Tables V. and VI. assume that both with regard to fathers and children this is the case. If this assumption be made then the table must be a symmetrical one, and as one cannot get a symmetrical table by combining the two sources of information, one has to construct a separate table out of each. In Table V. the information from Table I. only is used, and in Table VI. that from Table II.

Tables VII., VIII., IX. and X. were made for mothers in exactly the same way as III., IV., V. and VI. were for the fathers. In Table XI. the values of the correlation coefficients for each of these eight tables are set down. It will be

TABLE VII.

Mothers.

Children.	Deaf	Hearing
Hearing ...	4632	7,722,294
Deaf... ...	469	13,121

$h = 2{\cdot}919020$
$H = {\cdot}005632$
$k = 3{\cdot}212070$
$K = {\cdot}002294$
$r = {\cdot}51$

TABLE VIII.

Mothers.

Children.		Deaf	Hearing
Hearing	...	4632	20,604,933
Deaf...	...	469	13,121

h =3·483478
H = ·000925
k =3·212070
K = ·002294
r = ·56

TABLE IX.

Mothers.

Children.		Deaf	Hearing
Hearing	...	4632	7,730,783
Deaf...	...	469	4632

$h = k$ =3·212070
$H = K$ = ·002294
r = ·62

TABLE X.

Mothers.

Children.		Deaf	Hearing
Hearing	...	3525	5,533,036
Deaf...	...	126	3525

$h = k$ =3·212070
$H = K$ = ·002294
r = ·45

TABLE XI.

Method of construction of Table	Value of r	
	Paternal	Maternal
Tables I. and II. both used, distribution of deafness among parents assumed to be that of general population	0·51	0·51
Tables I. and II. both used, distribution of deafness among children assumed to be that of general population	0·57	0·56
Table I. only used, distribution of deafness both among parents and children assumed to be that of general population	0·62	0·62
Table II. only used, distribution of deafness both among parents and children assumed to be that of general population	0·46	0·45
Mean value of r	0·54	0·535

seen that the values of r vary between 0·45 and 0·62, with a mean value for paternal correlation of 0·54, and for maternal of 0·535.

Now we have two sources of inaccuracy, the tendencies of which are known, in arriving at these values. Firstly, the inclusion of persons whose deafness was due to causes acting on them from without. Although these are probably far less numerous than the whole number entered by Fay as "acquired" or "adventitious" cases, yet there are certainly enough of them present to considerably reduce the value of r in all the tables if, as it is generally supposed, this form of deafness is not inherited; it is however quite possible that the physical condition, which renders the owner of it liable to become deaf in this way, is inherited. Secondly, the percentage of deaf-mutes given in the census returns is in all probability too low. If a higher percentage had been allowed for in making the tables the effect would have been to reduce the value of r; as it was not allowed for one must consider that the latter error has to a certain extent balanced the former, though exactly to what extent is quite uncertain.

Another factor to be taken into account in considering the paternal and maternal correlation coefficients is that of "Assortative Mating" of the parents. The tendency for the deaf to marry the deaf is an exceedingly strong one, and that for obvious reasons which need not be dwelt on here. In Table I. a total of 2954 marriages is recorded, and in 2373 of these both husband and wife were deaf, in 344 the husband was deaf and the wife hearing, and in 237 the reverse was the case.

In order to determine exactly the coefficient of assortative mating, that is to say the correlation between husband and wife, I constructed two correlation tables from these data; one of these assumed that the distribution of deafness among husbands was normal, the other that this was the case for wives. Owing to the very large value of r in both cases, it was not found possible to determine it exactly, as this would have involved the solution of an equation containing very much higher powers of x than x to the twentieth. The tables are therefore not included here, but the value of r is certainly considerably higher than ·9.

(2) *Fraternal Correlation.*

As has been mentioned previously Fay has endeavoured to give in his tables the size of the families in which the 4471 pairs of husband and wife are contained, and also the number of deaf-mutes contained in them. In his statement of this, brothers are not distinguished from sisters, but all are included and numbered together. In several cases the numbers are not known or not stated definitely. Excluding these we find that the remaining deaf husbands and wives are included in 3837 families, the distribution of which with regard to size is shown in the last column of Table XII. It is also shown graphically in Figure I., and is there fitted with the skew curve (Prof. Pearson's General Probability Curve of Type I.), which seems to represent the observations with fair accuracy.

Of each of these families every possible pair of deaf and deaf, of deaf and hearing, and of hearing and hearing brothers and sisters was made. To take a

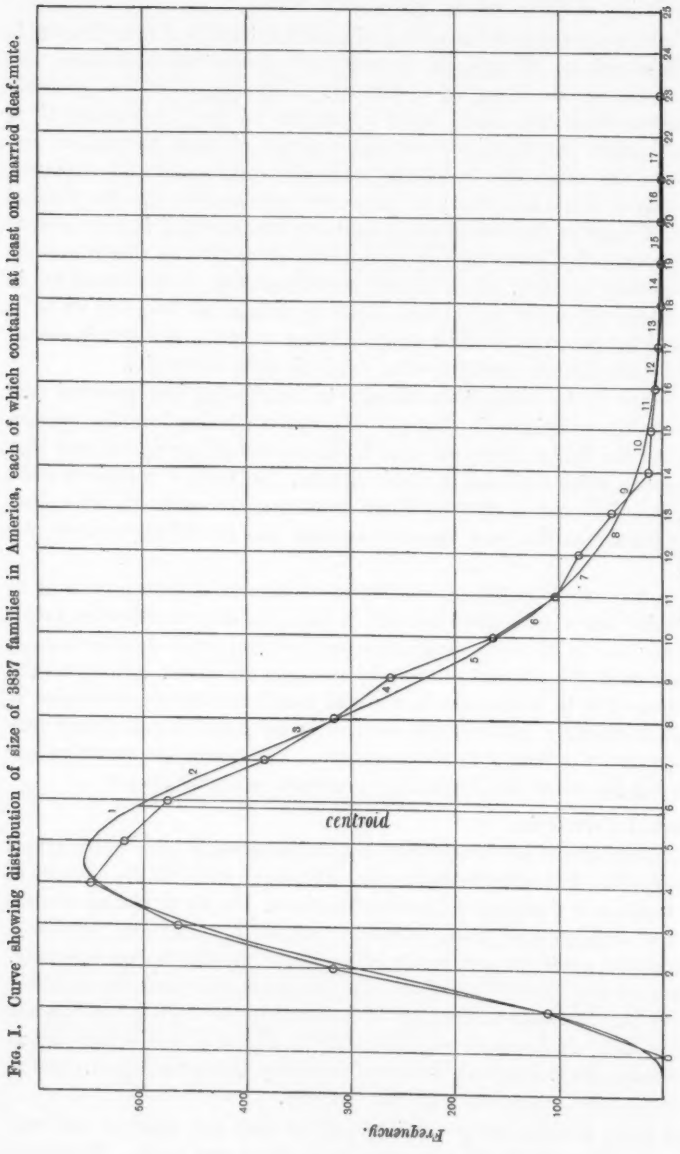


FIG. I. Curve showing distribution of size of 3837 families in America, each of which contains at least one married deaf-mute.

Equation for Curve of Type I.: $y = 555{\cdot}79\left(1+\frac{x}{61{\cdot}9957}\right)^{39{\cdot}9766}\left(1-\frac{x}{5{\cdot}2895}\right)^{3{\cdot}4108}$.

Mean size of Family = 5·8328.

Origin of curve at 4·5817.

TABLE XII.

Comparing distribution of size of families which contain at least one congenitally deaf person, with that of those containing at least one person said to have acquired deafness.

Number in Family	Nature of Deafness			Totals
	Congenital	Acquired	Doubtful	
1	32	73	6	111
2	83	218	18	317
3	118	337	17	466
4	180	365	16	550
5	168	344	17	518
6	144	323	18	476
7	117	260	16	383
8	97	215	12	316
9	88	176	11	262
10	58	103	8	163
11	46	55	5	104
12	31	50	1	81
13	18	35	—	49
14	5	11	—	14
15	5	6	—	11
16	1	5	1	7
17	1	3	—	4
18	—	1	—	1
19	—	1	—	1
20	—	1	—	1
21	—	1	—	1
22	—	—	—	—
23	1	—	—	1
Totals	1193	2583	146	3837

N.B. 85 families are included both in the congenital and in the acquired class, as they contain at least one member of each of these two classes.

concrete example of this; consider a family of 12 persons containing 6 deaf members and 6 hearing. Taking the deaf members first each of them can be paired with each of the other deaf ones, thus making 30 pairs of deaf and deaf; they can also be paired with the hearing members, making 36 pairs of deaf and hearing; similarly the 6 hearing persons can be paired with the 6 deaf, making 36 pairs of hearing and deaf, and with one another, making 30 pairs of hearing and hearing. The results can be written in tabular form as in Table XIII. Out of a family of 12, 132 pairs can be made, and, speaking generally, out of a family of n members $n(n-1)$ pairs can be made. Table XIV. is the sum of tables such as XIII. made for each family.

Now there are altogether 5423 deaf persons in the families used, and the distinctive point of these families is that they contain at least one married deaf

TABLE XIII.

Showing pairs which can be made out of a family of 12 *persons,* 6 *deaf and* 6 *hearing.*

	Deaf	Hearing
Hearing ...	36	30
Deaf... ...	30	36

TABLE XIV.

Showing total number of pairs made out of the 3837 *families used.*

	Deaf	Hearing
Hearing ...	23,772	88,572
Deaf... ...	5226	23,772

member. They should therefore be compared with hearing people contained in families which have at least one married member. If they bear the same numerical relation to these that deaf people in general do to hearing people in general, then they occur in a total number of 5423 multiplied by 1,000,000/659 persons. Assuming that the mean size of families, which contain a married member but do not contain a deaf one, is the same as that of families which contain a deaf married one, and that the distribution of size is the same; then, as from 5423 deaf persons 23,772 plus 5226 (equals 28,998) pairs can be made with their brothers and sisters, the total number of pairs which can be made from $\frac{5423}{659} \times 1{,}000{,}000$ persons is $\frac{5423}{659} \times 1{,}000{,}000 \times \frac{28998}{5423} = 44{,}003{,}035$. Deducting from this number all the pairs that contain deaf-mutes there are left 43,950,262 pairs of hearing and hearing brothers and sisters, and by substituting this number for the number of pairs of hearing and hearing in Table XIV., Table XV. is obtained.

TABLE XV.

Fraternal Correlation Table constructed from Table XIV.

	Deaf	Hearing
Hearing ...	23,772	43,950,262
Deaf... ...	5226	23,772

$h = k = 3{\cdot}212070$
$H = K = {\cdot}002294$
$r = {\cdot}74$

With regard to the assumption that families which do not contain a deaf member are the same in point of size as those that do, I endeavoured to obtain evidence from the data at my disposal. It seems quite clear that, at any rate, a considerable proportion of those persons who are said to have acquired deafness must in reality have no sort of inborn tendency in this direction; the families which contain them can therefore be taken as more nearly approaching a sample of families whose only common character is that they contain at least one married member, than those which contain a congenital deaf-mute. If then it is found that the mean size and the variation in size in both of these classes is approximately the same, it could be argued that this is to a certain extent evidence that the assumption is a reasonable one. The families were accordingly divided into these two classes, and the first two columns in Table XII. show the distribution of size for each class. Table XVI. gives a comparison of the mean size of family, and of

TABLE XVI.

Showing difference in mean size of families containing at least one congenital deaf-mute and of those containing at least one member who has acquired deafness; also the difference in the standard deviations.

	Mean	Standard Deviation
Congenital ...	6·0821 ± ·0583	2·9862 ± ·0412
Acquired ...	5·7704 ± ·0386	2·9070 ± ·0273
Difference ...	·3117 ± ·0699	·0792 ± ·0495

the standard deviation. It will be seen that the mean size of families containing at least one congenital deaf-mute is larger than those containing one of the "acquired" variety by an amount which is rather more than four times as great as its probable error. There is thus a certain amount of reason for believing that families containing a deaf-mute member are in reality somewhat larger than those that do not. If this had been allowed for in making the correlation table the number of pairs of hearing and hearing brothers and sisters would not have been so great, and there would have been a consequent reduction in the value of r. The difference in the value of the standard deviation is not twice as great as its probable error, and is therefore probably not significant.

The value of r for fraternal correlation was found from Table XV. to be approximately ·74, which is a rather higher value than is normal, though not very much higher than that for the amount of red in coats of greyhounds of the same litter, which was found to be ·700 (*Biometrika*, Vol. II. p. 390).

Table XVII. explains itself. The mean number of deaf members per family in each case is not insisted upon, owing to the probable gross inaccuracies of the classification into cases of congenital and acquired deafness.

TABLE XVII.

Showing total number of deaf persons in families containing at least one congenitally deaf person, also in those containing at least one person who has acquired deafness.

Number of Deaf Persons in Family	Nature of Deafness			Totals
	Congenital	Acquired	Doubtful	
1	524	2260	101	2885
2	354	221	22	559
3	189	73	13	242
4	82	22	5	99
5	25	4	4	31
6	8	1	—	9
7	8	1	1	9
8	1	1	—	1
9	2	—	—	2
Totals	1193	2583	146	3837

N.B. 85 families are included both in the congenital and in the acquired class, as they contain at least one member in each of these two classes.

Conclusions.

In conclusion it need hardly be pointed out that the results obtained can only be regarded as provisional, as the data on which they are based were not collected with a view to the construction of correlation tables, and are therefore not particularly well suited to this purpose.

With this reservation it may, however, be asserted that striking confirmation has been obtained of previous work on widely different characters, at any rate with regard to the correlation between father and children, and mother and children.

Quite recently our knowledge of this subject has been greatly enlarged and lucidly summarized by Prof. Pearson and Dr Lee in their paper, "On the Laws of Inheritance in Man" (*Biometrika*, Vol. II. p. 355); and it will be unnecessary to repeat here the excellent comparative tables that are given there. It will be sufficient to mention one or two examples to show for what a broad range of characters parental inheritance has been measured, and to quote the general conclusion of the authors on this point.

While it has been found in the present paper that the mean of the results obtained by four different methods for the inheritance of deafness is ·54 for fathers and children and ·535 for mothers and children (*vide* Table XI.), the mean value of parental inheritance of stature in man is ·506, of span, ·459, and of eye colour, ·495; similarly that of coat-colour in horses is ·522, in Basset hounds ·524, and in

greyhounds ·507 (*Biometrika*, Vol. II. p. 379, Table V.). The general conclusion is expressed as follows:—"The coefficient of parental heredity varies to some extent from character to character in man, having a mean value of about ·46. This value is however in quite close agreement with the results obtained for other species, and we may roughly say that parental heredity in the species hitherto dealt with is close to ·5." (*Biometrika*, Vol. II. p. 396 (V.).)

The value obtained for fraternal correlation is less satisfactory, but if it is not directly confirmatory of, it is in no way inconsistent with previous determinations for other characters, for the range of variation of fraternal correlation seems to be much greater than that of parental. Thus ·74 here arrived at is much larger than the mean value for other characters in man (·5), but it does not differ much from the values obtained for coat-colour in horses and dogs, or for certain characters in waterfleas and aphides, all of which will be found duly set forth in the paper referred to above.

One more conclusion can be drawn, and that is the urgent necessity for the collection of satisfactory data, on which to work at this and other problems in heredity. The Interdepartmental Committee on Physical Deterioration, in their report, published in 1904, strongly recommend that an Anthropometric Survey should be undertaken by the Government at the earliest possible moment (*vide* Section 54). Should this recommendation ultimately be adopted, there is no reason why the scope of this survey should not be somewhat extended so as to include the collection of data bearing on heredity; but till then one must make what use one can of any material that one can get.

Postscript. Since writing the above I have been able to obtain values for paternal, maternal, and fraternal correlation from information collected in England. For the first two the following material was used:

(1) A record of the offspring of marriages in which one or both parents were deaf from the register of the National Association in Aid of the Deaf and Dumb. Of 161 marriages there recorded both parents were deaf in 127, while in 26 the husband was deaf and the wife was either hearing or not stated to be deaf, and in 8 the reverse was the case. Thus there were in all 153 marriages in which the husband was deaf, of which 125 had proved fertile by the time of the information. The total number of children was 435, of which 52 occurring in 25 families were deaf. There were 135 marriages in which the wife was deaf. 109 of these had resulted in a total number of 379 children, of which 50 occurring in 25 families were deaf. The only part of this used in the tables was that in families with a deaf father there were 383 hearing children and 52 deaf, and in families with a deaf mother there were 329 hearing children and 50 deaf.

(2) The rest of the material used is contained in Table XVIII. The number of deaf parents given there is probably much understated. Both Dr Kerr, to whose kindness I owe the information concerning the L.C.C. schools, and Mr Moore, who sent me that relating to Margate, warned me of this.

TABLE XVIII.

Source of Information	Total Number of Children	With Father Deaf only	With Mother Deaf only	With both Parents Deaf
Birmingham. Royal Institution for the Instruction of Deaf and Dumb children	679	1	6	1
Exeter. Royal West of England Institution for the Deaf and Dumb	66	1	1	—
London. 10 L. C. C. Schools	413	1	—	4
Margate. Royal Asylum for the Deaf and Dumb ...	1302	—	—	27
Newcastle-on-Tyne. Northern Counties Institution for the Deaf and Dumb	154	—	2	4
Stoke-on-Trent. North Staffordshire Joint School Authority's Blind and Deaf School	170	2	1	2
Totals	2784	5	10	38

TABLE XIX.

Fathers.

Children.	Deaf	Hearing
Hearing ...	383	924,329
Deaf... ...	52	3315

h = 3·3087
H = ·001674
k = 2·6849
K = ·010913
r = ·45

TABLE XX.

Fathers.

Children.	Deaf	Hearing
Hearing ...	383	7,179,796
Deaf... ...	52	3315

h = 3·3087
H = ·001674
k = 3·8436
K = ·000247
r = ·58

Tables XIX. and XX. give the paternal correlation coefficient. They were constructed from the material described above in exactly the same way as Tables III. and IV. were from Fay's data, except that the rate per million of deaf persons used was that of the 1901 English census, namely 468·71. The mean value obtained from them is ·515, while that obtained from Tables III. and IV. is ·54.

Tables XXI. and XXII. correspond exactly with Tables VII. and VIII., and like them they give a mean value for maternal correlation of ·535.

The material from which the fraternal correlation coefficients were deduced is shown in Table XXIII. It consists of 459 families from the register of the

TABLE XXI.

Mothers.

Children.

	Deaf	Hearing
Hearing ...	329	805,373
Deaf... ...	50	2850

h =3·3087
H = ·001674
k =2·6887
K = ·010743
r = ·47

TABLE XXII.

Mothers.

Children.

	Deaf	Hearing
Hearing ...	329	6,183,965
Deaf... ...	50	2850

h =3·3087
H = ·001674
k =3·8408
K = ·000250
r = ·60

TABLE XXIII.

Showing Distribution of Deaf in Families used for Fraternal Correlation.

Number of Persons in each Family	National Association in Aid of the Deaf and Dumb									Exeter and Manchester				
	Number of Deaf in Family								Totals	Number of Deaf in Family				Totals
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*		*1*	*2*	*3*	*4*	
1	10	—	—	—	—	—	—	—	10	4	—	—	—	4
2	23	2	—	—	—	—	—	—	25	17	1	—	—	18
3	47	6	—	—	—	—	—	—	53	20	4	—	—	24
4	56	7	2	—	—	—	—	—	65	23	4	1	1	29
5	48	9	1	3	—	—	—	—	61	24	1	1	—	26
6	45	8	5	4	—	—	—	—	62	10	2	—	—	12
7	33	3	1	3	—	1	—	—	41	11	1	—	1	13
8	35	3	5	—	1	—	—	—	44	5	—	1	—	6
9	25	5	1	1	1	—	—	—	33	11	1	—	—	12
10	17	5	2	1	—	—	1	—	26	4	—	2	—	6
11	16	—	3	—	—	—	—	—	19	1	—	—	—	1
12	3	—	—	—	—	—	—	—	3	—	—	—	—	—
13	5	—	1	1	—	—	—	—	7	2	—	—	—	2
14	2	2	—	—	—	—	—	—	4	—	—	1	—	1
15	—	—	—	2	—	—	—	—	2	—	—	—	—	—
16	1	—	—	—	—	—	—	—	1	1	—	—	—	1
17	1	—	—	—	—	—	—	—	1	—	—	—	—	—
18	2	—	—	—	—	—	—	—	2	—	—	—	—	—
Totals	369	50	21	15	2	1	1	—	459	133	14	6	2	155

National Association : Mean size of family ... ... 6·1939 ± ·0954
Standard deviation ... ... 3·0299 ± ·0674
Mean number of deaf per family 1·3399

Exeter and Manchester : Mean size of family ... ... 5·1871 ± ·1489
Standard deviation ... ... 2·7493 ± ·1053
Mean number of deaf per family 1·2064

National Association, and a group of 155 families of which 64 came from the institution at Exeter and 91 from that at Manchester. From the 459 families Table XXIV. was made, and from the 155 Table XXV., both in exactly the same way as Table XV.

TABLE XXIV.

National Association in Aid of the Deaf and Dumb.

	Deaf	Hearing
Hearing ...	2880	7,243,407
Deaf... ...	518	2880

$h = k = 3{\cdot}3087$
$H = K = {\cdot}001674$
$r = {\cdot}72$

TABLE XXV.

Exeter and Manchester.

	Deaf	Hearing
Hearing ...	727	1,737,273
Deaf... ...	88	727

$h = k = 3{\cdot}3087$
$H = K = {\cdot}001674$
$r = {\cdot}68$

The values of r were found to be ·72 and ·68 respectively, whereas ·74 was given by Table XV. It may thus be said that with regard to parental and fraternal correlation the English material gives results which agree very well with the American, if one takes into consideration the roughness both of the material and of the methods employed.

The mean size of the families from the National Association Register is $6{\cdot}1939 \pm {\cdot}0954$; that of the Exeter and Manchester families is $5{\cdot}1871 \pm {\cdot}1489$. The reason for this difference is that in the former case the families are probably all complete, whereas in the latter they are certainly incomplete in many cases, the number being taken at the time of the entry of one of their deaf members, generally aged 8 or 9, into one of the two institutions.

ON THE CORRELATION BETWEEN VACCINATION AND SMALLPOX IN THE LONDON EPIDEMIC, 1901–2.

By F. M. TURNER, M.D.

At the suggestion of Prof. Pearson I have undertaken the calculation of the smallpox statistics of the Metropolitan Asylums Board Hospitals for the epidemic of 1901–2, in continuation of the two papers already published on the same subject by Macdonell in this Journal; and the results are given in this paper. I also include a discussion of several points, which have occurred to me in the course of the calculations, which seem to me to have considerable importance in altering the values of the coefficients. Of these the most important is a suggestion as to the probable frequency distribution of severity among smallpox cases, which I call a "curtailed" normal distribution, and a discussion is given of the mathematical formulae appropriate to such a distribution. All the data were obtained from the *Annual Reports of the Metropolitan Asylums Board* for the years 1901 and 1902.

Taking a fourfold division into vaccinated and unvaccinated and into recoveries and deaths, as in the above-mentioned papers, we get the following tables for the calculation of r. They only differ in the treatment adopted for the doubtful cases, which in Table I. are classed with the vaccinated, in Table III. with the unvaccinated, and in Table II. have been omitted altogether.

TABLE I.

All Cases, 1901—2. *Vaccination and Severity.*

	Recoveries	Deaths	Totals
Unvaccinated	1525	753	2278
Vaccinated or doubtful ...	6505	876	7381
Totals ...	8030	1629	9659

$r = -\cdot4246 \pm \cdot0049.$

TABLE II.

All Cases except those in which Vaccination was Doubtful, 1901—2. *Vaccination and Severity.*

	Recoveries	Deaths	Totals
Unvaccinated ...	1525	753	2278
Vaccinated ...	6240	705	6945
Totals ...	7765	1458	9223

$r = -\cdot 4741$.

TABLE III. *All Cases,* 1901—2. *Vaccination and Severity.*

	Recoveries	Deaths	Totals
Unvaccinated or doubtful ...	1790	924	2714
Vaccinated	6240	705	6945
Totals	8030	1629	9659

$r = -\cdot 4922$.

Of these three values I prefer that of Table I., as I believe that a majority of the doubtful class were really vaccinated. This class includes not only those cases stated to have been vaccinated, in whom the scar was certainly absent, but also those in which the presence or absence of the scar could not be affirmed with certainty on account of the abundance of the smallpox eruption. This latter class is much more numerous than the former and naturally has a high death-rate. Consequently the position assigned to the doubtful class in the table has an influence on the value of r out of proportion to the size of the class.

The reasons for believing the majority of the doubtful cases to have been vaccinated are:

(*a*) In those cases where direct observation fails, it seems reasonable to attach value to the statement that the patient had been vaccinated, as likely to be right in the majority of instances.

(*b*) Smallpox affects vaccinated and unvaccinated persons at very different ages; not only is the prevailing age different in the two groups, but also the two groups only overlap to a moderate extent. The age distribution among the doubtful cases coincides most closely with that of the vaccinated class, as is seen in the adjoining table.

TABLE IV.

Age incidence among the Vaccinated, Unvaccinated and Doubtfully Vaccinated.

Age	*0–9*	*10–19*	*20–29*	*30–39*	*40–49*	*50–59*	Over *60*	Totals
Vaccinated ...	134	1163	2517	1755	893	320	163	6945
Doubtful cases ...	33	73	98	77	62	52	41	436
Unvaccinated ...	1274	619	244	90	31	13	7	2278

(*c*) On the other side it is often argued that the death-rate among the doubtful class is high, and corresponds closely with that among the unvaccinated. Hence it is assumed that the majority of "doubtful" cases were probably unvaccinated. I do not agree with the conclusion, for the high death-rate (higher than among the admittedly unvaccinated) may be due to the fact that in cases with very severe eruption it is impossible to be sure that a scar is present.

It is noticeable that the above values are lower than any of those found by Macdonell for previous epidemics, as well as for those he obtained for the same epidemic in London. His figures for the latter were obtained before the *M. A. B. Report* for 1902 was published, and are subject to the drawback of including an undue proportion of deaths. I cannot explain this better than by quoting the article in the *Times* from which his figures were drawn. "Returns during the "progress of an outbreak have a serious drawback, which must be carefully borne "in mind. It is obvious that only completed cases can be dealt with, as the "result of those still under treatment is uncertain; but since death usually takes "place at a comparatively early stage of the illness, the fatal cases are always in "advance of the recoveries.....The case mortality therefore always appears very "much higher during an outbreak than it really is. In the present instance it "is probably more than twice as high as it should be."

The following two tables are copied from his paper:

TABLE A.

London, Cases admitted to M. A. B. Hospitals, and completed to Recovery or Death, up to November 30th, 1901*.

	Recoveries	Deaths	Totals
Unvaccinated ...	30	60	90
Vaccinated ...	195	45	240
Totals	225	105	330

$r = -\cdot6605 \pm \cdot0406$.

TABLE B.

London, Cases admitted to M. A. B. Hospitals, and completed to Recovery or Death, up to the end of 1901.

	Recoveries	Deaths	Totals
Unvaccinated ...	96	98	194
Vaccinated ...	652	108	760
Totals	748	206	954

$r = -\cdot5779 \pm \cdot0311$.

* I have altered the sign of *r*, and the arrangement of the table to make it comparable with my own tables.

From these two tables and a further table of Islington cases, brought up to the end of the epidemic, he concludes that the correlation diminished as the epidemic progressed.

I have calculated the correlation separately for the two years 1901 and 1902.

TABLE V.

London, all Cases admitted to M. A. B. Hospitals during the year 1901.

	Recoveries	Deaths	Totals
Unvaccinated	253	119	372
Vaccinated or doubtful ...	1198	173	1371
Totals	1451	292	1743

$r = -\cdot 3851$.

TABLE VI.

London, all Cases admitted to M. A. B. Hospitals during the year 1902.

	Recoveries	Deaths	Totals
Unvaccinated	1272	634	1906
Vaccinated or doubtful ...	5307	703	6010
Totals	6579	1337	7916

$r = -\cdot 4330$.

I therefore conclude that the correlation did not diminish with the progress of the epidemic. Table V. is strictly comparable with Table B in period, and differs from it chiefly in taking account of the ultimate result of all those cases remaining in hospital at the end of the year. Both my tables include the doubtful cases which are omitted in Macdonell's. The effect is to lower the values of r, but not as can be seen by comparing Tables I. and II. to an extent sufficient to account for the difference.

I have also calculated the coefficients of correlation between severity of disease and number, area and foveation of scars from the figures for the vaccinated cases only; severity of disease being estimated as in the previous tables by the fact of death or recovery.

TABLE VII.

Vaccinated cases only, 1901—2. *Area of Vaccination Scars and Severity.*

	Recoveries	Deaths	Totals
Area $< \frac{1}{2}$ inch ...	1402	293	1695
Area $> \frac{1}{2}$ inch ...	4784	379	5163
Totals ...	6186	672	6858

$r = -\cdot 2997$.

TABLE VIII.

Vaccinated cases only, 1901—2. *Number of Vaccination Scars and Severity.*

	Recoveries	Deaths	Totals
Less than 4 scars ...	3449	506	3955
4 scars or more ...	2761	169	2930
Totals	6210	675	6885

$r = -\cdot 2631$.

TABLE IX.

Vaccinated cases only, 1901—2. *Foveation of Scars and Severity.*

	Recoveries	Deaths	Totals
Scars less than half foveated ...	5203	602	5805
Scars more than half foveated...	980	58	1038
Totals	6183	660	6843

$r = -\cdot 1664$.

These figures agree well with those given by Macdonell for the Glasgow epidemic of 1900–1, less closely for the epidemic of 1892–5 in the same town. The slight differences in the total numbers observed in the three tables are due to the measurement of area or the counting of the number of scars having presented special difficulty in certain cases.

An important practical point to settle, if possible, is whether these three correlations are independent or two of them secondary to the third. It seems probable *à priori* that the protection afforded by vaccination is proportional to the area of the vaccine vesicles, which may be regarded as laboratories for the production of an antitoxin. If this were so the correlation between number of scars and severity would be secondary to that between number and area of scars.

TABLE X.

Vaccinated cases only, 1901—2. *Number and Area of Scars.*

	Area $< \frac{1}{2}$ inch	Area $> \frac{1}{2}$ inch	Totals
Less than 4 scars ...	1526	2398	3924
4 scars or more ...	165	2761	2926
Totals	1691	5159	6850

$r = +\cdot 6648$.

Denoting by r_{AS}, r_{NS}, r_{AN}, the total correlations between area of scars, number of scars and severity of disease taken in pairs, and ρ_{AS}, ρ_{NS}, the partial correlations

between area and number of scars with severity of disease, we have, when (i) area is constant and (ii) number is constant respectively,

$$r_{NS} = -\cdot2631, \quad r_{AS} = -\cdot2997, \quad r_{AN} = +\cdot6648.$$

$$\rho_{NS} = \frac{r_{NS} - r_{AS}r_{AN}}{\sqrt{(1 - r^2_{AS})(1 - r^2_{AN})}} = -\cdot0897,$$

$$\rho_{AS} = \frac{r_{AS} - r_{NS}r_{AN}}{\sqrt{(1 - r^2_{NS})(1 - r^2_{AN})}} = -\cdot1733.$$

Dealing in the same manner with area and foveation of scar, we have the following table:

TABLE XI.

Vaccinated cases only, 1901—2. *Area and Foveation of Scars.*

	Area $< \frac{1}{2}$ inch	Area $> \frac{1}{2}$ inch	Totals
Scars less than half foveated ...	1530	4255	5785
Scars more than half foveated...	145	893	1038
Totals	1675	5148	6823

$r = +\cdot2386.$

And $$r_{AS} = -\cdot2997, \quad r_{FS} = -\cdot1664, \quad r_{AF} = +\cdot2386,$$

$$\rho_{FS} = \frac{r_{FS} - r_{AS}r_{AF}}{\sqrt{(1 - r^2_{AS})(1 - r^2_{AF})}} = -\cdot1012,$$

$$\rho_{AS} = \frac{r_{AS} - r_{FS}r_{AF}}{\sqrt{(1 - r^2_{FS})(1 - r^2_{AF})}} = -\cdot2719.$$

It seems therefore not improbable that among cases with equal area of vaccination marks, there is no difference in the amount of protection afforded by different numbers of scars. The apparent protection afforded by foveation of the scars, though not very great in itself, can with less probability be attributed to secondary association of foveation with large scars.

It would be a good thing if these two points could be definitely settled, for a great deal of time is spent in smallpox hospitals in observing and recording these data, some of which may possibly be unnecessary.

The fact that ρ_{NS} as above calculated is not equal to 0 may be due to the fact that some scars grow considerably between infancy and adult life. This of course would tend to lessen r_{AS} (the area being measured many years after the operation), but would leave r_{NS} unaltered unless, as is probably not common, one or more scars disappeared entirely. The above formulae also apply only to linear correlation, while the correlation in these cases may be skew.

I have paid much attention to the subject of age in these statistics, as the subject seems to me important from two points of view. It is known that the protection afforded by vaccination against attack is not permanent, but diminishes with lapse of time. A quantitative estimate of the amount of protection soon after vaccination and after the lapse of different intervals of time would be a useful guide as to the proper period at which to perform revaccination. Since no statistics are available as to the protection against attack, it seems reasonable to enquire into the time-decrement of the protection against death and take this as a guide. The second reason for enquiry as to age of the patients is the heterogeneity of the figures given above as regards age. If we take the correlation coefficient between vaccination and severity to be a measure of the protection afforded by the operation, it seems reasonable to demand that the two groups, vaccinated and unvaccinated, should be similar in all respects, so far as that is possible of attainment. But this is far from the case in respect to age in the above tables. Table IV. shews that the ages of the vaccinated and unvaccinated differed greatly, and this would tend to alter the value of the correlation. Using the language of multiple correlation there is a very high correlation, among the patients treated in hospital, between age and vaccination. That being so, we cannot affirm *à priori* a close correspondence in magnitude, nor even in sign, between the total correlation r_{VS} between vaccination and severity, and the partial correlation ρ_{VS} between the same factors.

To ascertain the time-decrement of protection I have divided up the cases into five age groups and worked out the value of r in each. Nearly all the cases who were vaccinated at all were vaccinated in infancy, and the number of these who had been revaccinated was small. The assumption that the age of the patient measures the number of years that had elapsed between the last operation of vaccination and the attack of smallpox is in most cases correct.

The following shews the number of revaccinated cases admitted:

TABLE XII.

Number of Revaccinated Persons Admitted in 1901 *and* 1902.

Interval in Years since Revaccination.

	0—10	*10—20*	*20* and over	Uncertain	Totals
1901	12	24	45	1	82
1902	34	79	161	2	276
Totals	46	103	206	3	358

The figures for 1902 are not complete in that no record was kept of revaccination at Gore Farm Hospital. As the cases admitted to that hospital numbered 565 only, as against 7316 to the two hospitals where the record was kept, the omitted cases cannot have been numerous.

A complete list of the 276 cases recorded in 1902 is published in the report. From this I have compiled the following details as to age:

TABLE XIII.

Ages of Revaccinated Persons Admitted in 1902.

	Under *15*	*15—20*	*20—25*	*25—30*	*30—40*	*40—50*	*50—60*	Over *60*	Totals
No. of cases	3	—	13	21	72	103	44	20	276
No. of deaths	—	—	1	—	9	7	8	2	27

From these figures it will be seen that Tables XVI. and XVII. below only contain 3 revaccinated cases, and the high value of r at this period is not due to prevalent revaccination.

TABLE XIV.

Patients under 10 *years of age*, 1901—2.

	Recoveries	Deaths	Totals
Unvaccinated	832	442	1274
Vaccinated or doubtful ...	159	8	167
Totals	991	450	1441

$r = -\cdot5567$.

TABLE XV.

Patients under 10 *years of age*, 1901—2.

	Recoveries	Deaths	Totals
Unvaccinated or doubtful ...	859	448	1307
Vaccinated	132	2	134
Totals	991	450	1441

$r = -\cdot6636$.

TABLE XVI.

Patients from 10 *to* 19 *years of age*, 1901—2.

	Recoveries	Deaths	Totals
Unvaccinated	469	150	619
Vaccinated or doubtful ...	1198	38	1236
Totals	1667	188	1855

$r = -\cdot6414$.

TABLE XVII.

Patients from 10 *to* 19 *years of age,* 1901—2.

	Recoveries	Deaths	Totals
Unvaccinated or doubtful ...	527	165	692
Vaccinated	1140	23	1163
Totals	1667	188	1855

$r = -\cdot 6933$.

TABLE XVIII.

Patients from 20 *to* 29 *years of age,* 1901—2.

	Recoveries	Deaths	Totals
Unvaccinated	158	86	244
Vaccinated or doubtful ...	2428	187	2615
Totals	2586	273	2859

$r = -\cdot 5364$.

TABLE XIX.

Patients from 20 *to* 29 *years of age,* 1901—2.

	Recoveries	Deaths	Totals
Unvaccinated or doubtful ...	216	126	342
Vaccinated	2370	147	2517
Totals	2586	273	2859

$r = -\cdot 6212$.

TABLE XX.

Patients from 30 *to* 39 *years of age,* 1901—2.

	Recoveries	Deaths	Totals
Unvaccinated	49	41	90
Vaccinated or doubtful ...	1548	284	1832
Totals	1597	325	1922

$r = -\cdot 4093$.

TABLE XXI.

Patients from 30 *to* 39 *years of age,* 1901—2.

	Recoveries	Deaths	Totals
Unvaccinated or doubtful. ...	90	77	167
Vaccinated	1507	248	1755
Totals	1597	325	1922

$r = -\cdot 4790$.

TABLE XXII.

Patients 40 *years of age and over*, 1901—2.

	Recoveries	Deaths	Totals
Unvaccinated	17	34	51
Vaccinated or doubtful ...	1172	359	1531
Totals	1189	393	1582

$r = -\cdot 4771$.

TABLE XXIII.

Patients 40 *years of age and over*, 1901—2.

	Recoveries	Deaths	Totals
Unvaccinated or doubtful ...	98	108	206
Vaccinated	1091	285	1376
Totals	1189	393	1582

$r = -\cdot 4440$.

Of the two methods of dealing with the doubtful cases I have given my reasons above for preferring their inclusion with the vaccinated rather than the unvaccinated. As between Tables XIV. and XV., however, I think it possible that a majority of the "doubtful cases" under 10 were unvaccinated, in which case the value given by Table XV. would be the more accurate.

On comparing the values of either series, however, we find that the value of r does not vary very much in the three periods under 30 years of age, whereas after the age of 30 there is a moderate fall in value. The maximum occurs in the period from 10 to 20. This result is in striking contrast to the opinion of the members of the medical profession as to the duration of immunity against attack; which is that the protection afforded by vaccination is highest for the first few years after the operation and then falls off rapidly, so that at 10 or 12 years of age revaccination becomes desirable. So long as no statistics are available of the number of persons exposed to risk of infection by smallpox, it is impossible to calculate by mathematical methods the protection afforded by vaccination against attack. But the above opinion as to the duration of vaccination-immunity is really arrived at by methods identical in principle to the above fourfold tables, that is by recording the numbers attacked of vaccinated and unvaccinated persons at different ages, and contrasting these with the very roughly known proportion existing between vaccinated and unvaccinated persons in the community at large. Inside a smallpox hospital there is found a very high correlation between age and vaccination, say ·9 or ·8 (see Tables XXIV., XXVI. and XXVIII.). If there existed an equally high correlation among the whole community between age and vaccination, we might assume that vaccination either failed to protect, or protected with equal force throughout life. But if there is only slight correlation between

these two conditions among the population outside, which is most probably the case, the protection gained by the vaccinated in infancy must diminish as age advances.

We seem therefore driven to the conclusion that protection against attack and protection against death, though both produced by the same operation, disappear at widely different rates. In a later part of this paper, however, I put forward a suggestion which I think accounts for the discrepancy in the figures, and allows us to suppose that the protections against attack and against death produced by vaccination are the same, not only in cause, but also quantitatively as measured by correlation coefficients.

The effect of the age of the patients upon the value of r in Table I. is perhaps best shewn by comparing it with the values derived from Tables XIV.—XXIII.

The mean value of r from the five tables, in which the doubtful cases have been included with the vaccinated, is $-\cdot5234$, against $-\cdot4246$ of Table I. And the mean value from the other five tables is $-\cdot5795$ against $-\cdot4922$ of Table III. The effect has therefore been to diminish r. It is interesting, however, to use the formulae for multiple correlation to test the result.

Taking the division of age at 10 years, we have the following table:

TABLE XXIV.

All Cases, 1901—2*.

	Under *10*	Over *10*	Totals
Unvaccinated ...	1274	1004	2278
Vaccinated ...	134	6811	6945
Totals ...	1408	7815	9223

$r = +\cdot912$.

TABLE XXV.

All Cases, 1901—2.

	Recoveries	Deaths	Totals
Under *10* ...	991	450	1441
Over *10* ...	7039	1179	8218
Totals ...	8030	1629	9659

$r = -\cdot3150$.

* I have inadvertently omitted the doubtful cases in calculating this table. As these high correlations are laborious to work out I have not thought it necessary to recalculate the value with the doubtful cases included. In all the tables of age and vaccination I have carried the series in r to twelve terms, and the terms beyond would certainly influence the third figure, probably the second also. The Tables XIV. and XV. also give series in r which converge slowly. In these two cases also I have stopped at twelve terms.

From Tables I., XXIV. and XXV. we get

$$r_{VS} = -\cdot4246, \quad r_{AV} = +\cdot912, \quad r_{AS} = -\cdot3150,$$

$$\rho_{VS} = \frac{r_{VS} - r_{AV}r_{AS}}{\sqrt{(1 - r^2_{AV})(1 - r^2_{AS})}} = -\cdot3530.$$

If we take the age division at 20 or 30 years we get the four Tables XXVI.—XXIX.

TABLE XXVI.

All Cases, 1901—2.

	Under *20*	Over *20*	Totals
Unvaccinated...	1893	385	2278
Vaccinated or doubtful ...	1403	5978	7381
Totals ...	3296	6363	9659

$r = +\cdot821.$

TABLE XXVII.

All Cases, 1901—2.

	Recoveries	Deaths	Totals
Under *20* ...	2658	638	3296
Over *20* ...	5372	991	6363
Totals ...	8030	1629	9659

$r = -\cdot0904.$

Whence we deduce

$$r_{VS} = -\cdot4246, \quad r_{AV} = +\cdot821, \quad r_{AS} = -\cdot0904,$$

$$\rho_{VS} = -\cdot6178.$$

TABLE XXVIII.

All Cases, 1901—2.

	Under *30*	Over *30*	Totals
Unvaccinated	2137	141	2278
Vaccinated or doubtful ...	4018	3363	7381
Totals ...	6155	3504	9659

$r = +\cdot671.$

TABLE XXIX. *All Cases*, 1901—2.

	Recoveries	Deaths	Totals
Under *30* ...	5244	911	6155
Over *30* ...	2786	718	3504
Totals ...	8030	1629	9659

$r = +\cdot1360.$

Whence we deduce

$$r_{VS} = -\cdot4246, \quad r_{AV} = +\cdot671, \quad r_{AS} = +\cdot1360, \quad \rho_{VS} = -\cdot7022.$$

The discrepancy between the above three values of ρ_{VS} is largely due to the want of linear correlation between age and severity of disease as measured by death or recovery. Another possible cause has occurred to me, which may have contributed to this result; and as this cause would also account for the small value shewn above for the diminution of vaccination protection with time, I am inclined to lay stress upon it.

Let us take any individual out of a population exposed to smallpox. Three things may happen to him, either he may escape the disease altogether, he may contract smallpox in a mild form and eventually recover, or he may contract it in a severe form and die. Into which of the three classes any individual falls, depends on his power of resistance to the disease or his immunity, and on the amount of virulence of the infection to which he is exposed. In some diseases, as diphtheria and tetanus, we are acquainted with immunity produced by a chemical substance, and which may therefore be quantitatively measured. In diphtheria and tetanus the above three classes correspond to definite quantitative amounts of antitoxin. In smallpox we have not this knowledge, but the analogy makes it justifiable to regard immunity as a quantity, though not as yet measurable. The immunity may be natural, or acquired by vaccination, or the result of the two causes combined.

If in such a population we desire to examine the correlation between severity of disease, which is the observed fact rendering the immunity visible to us*, with any other factor as vaccination, age, &c., we may make a sixfold table, thus:

TABLE C.

State as regards Disease.

State as regards Vaccination.	No Disease	Disease and Recovery	Disease and Death	Totals
Unvaccinated ...	*a*	*b*	*c*	*s*
Vaccinated ...	*d*	*e*	*f*	*t*
Totals ...	*g*	*l*	*m*	*N*

* The influence of exposure to different degrees of virulence of infection is probably also an important factor, but is ignored, because at present beyond our observation. In dealing with a large population there will probably be no marked differences in this respect between the vaccinated and the unvaccinated classes.

where $$g = a + d, \quad l = b + e, \quad m = c + f,$$
$$s = a + b + c, \quad t = d + e + f,$$
$$N = s + t = g + l + m.$$

This sixfold table can be turned into a fourfold table by adding the middle column to either the first or second. If the distribution of the two characters among the population is a normal one, the same value of r will be obtained by applying the "fourfold table" formulae to either of these. This, however, will not be the same as the value calculated by the same formulae from the fourfold table obtained by omission of the first column, thus:

TABLE D.

	Recovered	Died	Totals
Unvaccinated ...	b	c	u
Vaccinated ...	e	f	v
Totals ...	l	m	n

where $$u = b + c, \quad v = e + f, \quad n = l + m = u + v.$$

We have no accurate evidence on the distribution of immunity among the population at large, but the following reasons point to the assumption of normal distribution being more correct for the former sixfold than the latter fourfold table.

(1) The same agent, vaccination, is known to protect against attack and also against death. On the former supposition this fact is intelligibly represented in the table, for if an array representing a group of persons all vaccinated to an equal degree were to be vaccinated further, the severity of disease in each individual case would be reduced. It follows that some who would have contracted a fatal attack would now recover, i.e. some cases would be transferred from column 3 to column 2. Similarly others would be transferred from column 2 to column 1, and would not suffer from smallpox after exposure. In the fourfold table the first transfer is similarly represented, but the second is not. Whatever unit may ultimately be used to measure severity, the hypothesis of normal distribution applied to the latter table cannot explain a reduction of severity to 0 or a negative quantity.

(2) Among vaccinated persons the severe cases are much less frequent than the mild. The hypothesis of normal distribution amongst those affected with smallpox requires that in the total, and also in each array, the cases shall cluster round a certain mean severity and that cases of very great or very slight severity shall be extremely rare. If, however, normal distribution of immunity applies to the whole population, we can readily see how grades of slight severity can be the commonest.

Thus in Fig. 1 the abscissae represent degrees of immunity, which in the case of those who contract the infection are represented by the degree of severity

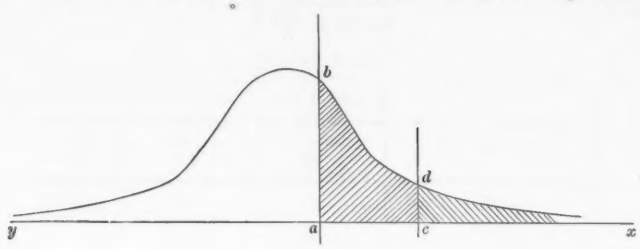


FIG. 1.

of the resulting disease. In those who escape infection the degree of immunity is greater than that required to prevent the disease appearing. It may be much more than this, but the excess is latent, and we have no means of observing it. The ordinates represent frequencies, the line *ab* divides those who escape infection from those who do not; the line *cd* divides the deaths from the recoveries.

The area *abdc* represents the number of recovered cases, and *xcd* the number of deaths. It is obvious that, if those who escape infection exceed half of those who are exposed, the frequencies of the different degrees of severity diminish without exception, as the degree increases. Probably the actual frequency curve among the vaccinated is of this type.

Assuming for the above reasons that the normal distribution applies to Table C we see that the same cannot hold also for Table D, though all the above calculations have been based on the latter assumption. It is therefore important to examine what is the effect of neglecting one column in a sixfold normal table and treating the remaining figures as if they conformed to a normal distribution.

Such a table I shall call a "curtailed" normal distribution, and the value of r obtained from the full table I shall call the true value. That which is obtained from the curtailed table by the use of any formula I shall call the "apparent" value and shall denote by r'.

As a first example I take the table on p. 214 of *Biometrika*, Vol. I., which is shewn by Macdonell to be a close fit to the normal correlation table. Drawing a line under "head breadth 15·0" I have kept all figures above the line and discarded all those below. The resulting table is therefore a curtailed correlation table, analogous to our smallpox tables drawn from the inmates of a smallpox hospital, all those who remained unaffected by the disease having escaped observation. A curtailed table in measurable characters would result if we collected statistics of height from soldiers, a population from whom all below a certain standard height had been already rejected.

Using product moments, I have found in the above case $r' = \cdot 2446$, whereas the true value $r = \cdot 4016 \pm \cdot 010$.

Using fourfold tables the value of r' varies very much according to the point chosen for division. Thus I have calculated the following three tables:

TABLE XXX.

Head Length.

Head Breadth.	*19·0* or less	Over *19·0*	Totals
Less than *15·0* ...	701	599	1300
15·0	109	119	228
Totals ...	810	718	1528

$r' = \cdot 0837$.

TABLE XXXI.

Head Length.

Head Breadth.	*19·0* or less	Over *19·0*	Totals
Less than *14·9* ...	619	458	1077
14·9 or *15·0* ...	191	260	451
Totals ...	810	718	1528

$r' = \cdot 2266$.

TABLE XXXII.

Head Length.

Head Breadth.	*19·0* or less	Over *19·0*	Totals
Less than *14·8* ...	526	348	874
14·8 to *15·0* ...	284	370	654
Totals ...	810	718	1528

$r' = \cdot 2595$.

It is seen that in every case r' is less than r.

If we use product moments a mathematical formula can be obtained.

Let
$$z = \frac{N}{2\pi\sqrt{1-r^2}\,\sigma_x\sigma_y} \cdot e^{-\frac{1}{2(1-r^2)}\left(\frac{x^2}{\sigma_x^2} - \frac{2rxy}{\sigma_x\sigma_y} + \frac{y^2}{\sigma_y^2}\right)} \quad \text{.....(I.)}$$

represent the equation to the correlation surface, and $x = h' = h\sigma_x$ be the equation to the bounding line.

Let $\bar{x}$, $\bar{y}$ be the coordinates of the centroid, r' the correlation, both the centroid and the correlation applying to the curtailed surface only.

Then
$$r' = \frac{\int_h^\infty \int_{-\infty}^{+\infty} (x-\bar{x})(y-\bar{y})\,z\,dx\,dy}{\left\{\int_h^\infty \int_{-\infty}^{+\infty} (x-\bar{x})^2\,z\,dx\,dy \times \int_h^\infty \int_{-\infty}^{+} (y-\bar{y})^2\,z\,dx\,dy\right\}^{\frac{1}{2}}} \quad \text{.....(II.)}$$

and $$\bar{x} = \frac{\int_h^\infty \int_{-\infty}^{+\infty} zx\,dx\,dy}{\int_h^\infty \int_{-\infty}^{+\infty} z\,dx\,dy} \quad \text{.....(III.).}$$

Since the centroid of every y array lies on the regression line, the centroid of the figure must also lie upon this line. Consequently

$$\bar{y} = \bar{x} \,.\, r \,.\, \frac{\sigma_y}{\sigma_x} \quad \text{.....(IV.).}$$

All the integrals in (II.) and (III.) can be reduced to the probability integral or directly integrated.

Putting $e^{-\frac{1}{2}h'^2} = \sqrt{2\pi}p$ and $\int_{h'}^\infty e^{-\frac{1}{2}x^2}\,dx = \sqrt{2\pi}q$,

we have $$\int_h^\infty \int_{-\infty}^{+\infty} z\,dx\,dy = \frac{N}{\sqrt{2\pi}\,\sigma_x} \int_h^\infty e^{-\frac{1}{2}\frac{x^2}{\sigma_x^2}}\,dx$$
$$= Nq,$$

and $$\int_h^\infty \int_{-\infty}^{+\infty} xz\,dx\,dy = \frac{N}{\sqrt{2\pi}\,\sigma_x} \int_h^\infty e^{-\frac{1}{2}\frac{x^2}{\sigma_x^2}} x\,dx$$
$$= Np \,.\, \sigma_x,$$
$$\therefore \bar{x} = \frac{p}{q}\sigma_x \text{..... .....(V.).}$$

From (IV.) $$\bar{y} = r \,.\, \frac{p}{q}\sigma_y \text{.....(VI.).}$$

Substituting the values from (V.) and (VI.) we obtain the following values for the three integrals in (II.):

$$\int_h^\infty \int_{-\infty}^{+\infty} (x - \bar{x})^2 z\,dx\,dy = N\sigma_x^2 \frac{hpq + q^2 - p^2}{q},$$
$$\int_h^\infty \int_{-\infty}^{+\infty} (y - \bar{y})^2 z\,dx\,dy = N\sigma_y^2 \frac{rhpq - r^2p^2 + q^2}{q},$$
$$\int_h^\infty \int_{-\infty}^{+\infty} (x - \bar{x})(y - \bar{y})\,z\,dx\,dy = Nr\sigma_x\sigma_y \frac{hpq + q^2 - p^2}{q}.$$

Substituting these values in (II.) we obtain:

$$r'^2 = \frac{r^2(hpq + q^2 - p^2)}{r^2(hpq - p^2) + q^2} \text{.....(VII.),}$$

or $$r^2 = \frac{r'^2 q^2}{q^2 + (1 - r'^2)(hpq - p^2)} \text{.....(VIII.).}$$

From (V.) we see that the distance between the centroids of the curtailed and the whole figures, resolved along the axis of x, is $\frac{p}{q}\sigma_x$. The distance of the limit of the curtailed figure from the latter centroid resolved in the same way is $h' = h\sigma_x$; so that $\frac{p}{q} - h$ is the expression for the distance of the curtailed centroid from the limiting straight line. It follows that $\frac{p}{q} - h$ is always positive.

Since p and q are essentially positive and $r^2 < 1$ we see from (VIII.) that it is impossible for r' to be numerically greater than r. If, however, $r = 0$ or 1, then r' also $= 0$ or 1.

The same can be shewn by differentiating $p - h$.

The equation (VIII.) can be written thus:

$$r^2 = \frac{r'^2}{1 + (1 - r'^2)\dfrac{hpq - p^2}{q^2}} \qquad \text{.....(IX.).}$$

In the case of smallpox correlations r' is the calculated apparent correlation. r is unknown, and cannot be calculated since h is also unknown. But it will be seen that when r' is constant r is greater when

$$\frac{p^2 - hpq}{q^2} = \left(\frac{p}{q}\right)^2 - h\left(\frac{p}{q}\right)$$

is less.

Calling this quantity z, the following gives values of z for a range of values of h:

h	z	h	z
−2·0	+·119	+0·1	+·658
1·5	·227	0·2	·662
1·0	·371	0·3	·700
0·5	·514	0·4	·715
0·4	·540	0·5	·729
0·3	·565	1·0	·796
0·2	·592	1·5	·848
0·1	·613	2·0	·854
−0·0	+·637	+3·0	1·30

It will be seen that, as h increases, z and consequently the discrepancy between r and r' also increase*.

* A simple geometrical interpretation of the quantity z is obtained thus. Let C, Fig. 2, represent the centroid of the whole figure, CX, CY the axes, PN the dividing line, and O the centroid of the curtailed figure. Then $CN = h\sigma_x$. Draw OM perpendicular to CX.

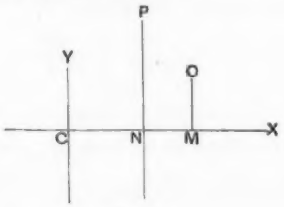


Fig. 2.

$$CM = \frac{p}{q} \cdot \sigma_x,$$

$$z = \left(\frac{p}{q}\right)^2 - h\left(\frac{p}{q}\right),$$

$$\therefore z \times \sigma_x^2 = CM^2 - CM \,.\, CN = CM \,.\, MN.$$

In the numerical example worked out above

$$r = \cdot 4016 \pm \cdot 010,$$

$$q = \frac{1528}{3000} = \cdot 509\,333$$

from the tables, $h = -\cdot 023\,398,$

$$p = +\cdot 399\,0257,$$

whence $r' = \cdot 2570.$

This agrees closely enough with the value ·2446 obtained by direct calculation considering the probable error of r.

In a curtailed normal distribution the use of fourfold tables and formulae is not strictly applicable, i.e. the result obtained from them is not the true correlation of the figure, for the formulae only apply to truly normal figures. Since, however, the formulae give a numerical result to whatever fourfold table they are applied, it is useful to discover the relation between the value so obtained, and the correlation coefficient of the whole figure.

I have not succeeded in finding a solution of the general problem. I have, however, calculated a fair number of curtailed normal tables by the fourfold formulae either chosen from the tables in Macdonell's paper on "Criminal Anthropometry," or formed by first using the fourfold formulae to get the frequency of a normal table with the value of $r = \cdot 5$ at various points of division; and then using these frequencies to form curtailed tables and calculating r'. It is rash to generalise from a few instances, but I think the following statements are true.

(1) The apparent correlation r' is always less than the true correlation r. The difference increases as the unobserved portion g increases, or as h decreases.

(2) In the same distribution, when r and h are constant, r' varies with the point of division. If this is close to the bounding line r' becomes very small.

An interesting problem, which may prove of practical application, is to solve a table of the form of Table C., when a and d are unknown, but $g = a + d$ is known. This may be done by the tedious process of trial and error. Give any value to a and from the sixfold tables deduce two fourfold tables. If the normal distribution holds good, the values of r obtained from the two tables will be identical.

As an example let us suppose that the number of persons exposed to infection in London was double the number who developed the disease. From this we get the following sixfold table.

TABLE E.

All Exposed to Infection.

	Not Diseased	Affected with Smallpox		Totals
		Recovered	Died	
Unvaccinated ...	a	1525	753	
Vaccinated ...	d	6505	876	
Totals ...	9659	8030	1629	19318

If $a = 450$ we get the two following fourfold tables:

	Not Diseased	Diseased	Totals
Unvaccinated ...	450	2278	2728
Vaccinated ...	9209	7381	16590
Totals ...	9659	9659	19318

$r_1 = -\cdot5373.$

	Alive	Dead	Totals
Unvaccinated ...	1975	753	2728
Vaccinated ...	15714	876	16590
Totals ...	17689	1629	19318

$r_2 = -\cdot5429.$

If $a = 400$ we find

$$r_1 = -\cdot5611, \qquad r_2 = -\cdot5496.$$

Whence by interpolation we get the correct solution:

$$a = 434, \qquad d = 9225,$$
$$r = -\cdot5451.$$

If, however, the number exposed to risk is not much larger than the number who contract the disease the values of r and r' more nearly coincide. Thus if the number exposed who do not contract the disease is equal to only one-tenth of those who do, we have $g = 966$, and I have calculated $a = 39$, $r = -\cdot4492$. On the other hand if the number of persons exposed to infection was several times larger than that of those who developed the disease, the true value of r considerably exceeded $-\cdot54$.

TABLE F.

	Not Diseased	Diseased		Totals
		Recoveries	Deaths	
Unvaccinated ...	39	1525	753	2317
Vaccinated ...	927	6505	876	8308
Totals ...	966	8030	1629	10625

$r = -\cdot 4492.$

I think the above figures afford some explanation of the apparent small diminution of vaccination protection as age advances. If my contention be correct that a normal distribution in immunity may more correctly be assumed for the whole population exposed to attack, than for that part only which actually contracts the disease, and appears in smallpox hospital statistics, then the values given above (Tables II. to IV. and XIV.—XXIII.) are apparent values of r corresponding to curtailed distributions.

If any time-diminution occurs, then the number of persons sufficiently protected to escape the disease altogether will be much larger at the ages 0–10 than at higher ages, say 20—30. Hence, using the above notation, in Table XIV. the value of g is much greater than in Table XIX. Hence $r - r'$ will be greater in the former case than in the latter. Although, therefore, the values of r' in the two cases are nearly equal, the values of r probably differ considerably.

The above investigation of curtailed distributions also gives a possible explanation of the small correlation between severity of disease and number, size and foveation [i.e. depth] of the scars.

For we have seen that if in any normal distribution, if one of the factors is divided into three classes, then leaving out either of the extreme classes from consideration will reduce the sixfold to a fourfold table. The resulting apparent correlation, however, will be less than the true value. Now in arranging our Tables I.—III. we have recognised only two degrees of vaccination, "vaccinated" and "unvaccinated." But if we split the vaccinated group into two, "four scars or more" and "less than four scars," we may arrange the whole as a sixfold table. Hence Table VIII. is seen to be a curtailed form of Table I.; and similar arguments apply to Tables VII. and IX. Macdonell, who discovered the lower correlation in these instances, remarks*: "It is obvious that in dealing with the last four tables we have descended to a much lower plane of correlation, and the results may possibly somewhat modify medical opinion as to the degree of significance of foveation, number of scars and scar area." If, however, the normal distribution held good for Table I., then a "lower plane of correlation" would necessarily follow in Tables VII.—IX., not as a medical, but as a mathematical phenomenon.

* *Biometrika*, Vol. I. p. 383.

Summary.

1. The correlation coefficient between severity of disease and previous vaccination in the late London epidemic was lower than in any of the previously investigated cases.

2. There seem stronger reasons for assuming normal distribution for the whole population exposed to risk of infection, than for the population attacked by the disease.

3. If the former assumption is correct the smallpox correlation tables of Pearson, Macdonell, and of this paper, correspond only to a portion of the normal correlation surface.

4. Values of r calculated from such "curtailed" normal distributions are lower than the true value, whether product moments are used or fourfold formulae, or contingency*. As to the first method a complete investigation is given, as to the other two I have failed to find a mathematical proof, and base the statement upon calculated instances, unfortunately not numerous.

5. Values of r have been calculated for each age group from the figures of the London epidemic. These appear to shew that the protection of vaccination lasts without perceptible diminution for about 30 years. This result is opposed to the opinion of most of those who have had experience of smallpox. My hypothesis given under (2) would explain the discrepancy, for the real values of the correlation at early ages would be considerably higher than the apparent values drawn from smallpox cases only.

6. The correlations between area of scars and severity, number of scars and severity, foveation of scars and severity in the London epidemic agree closely with those calculated by Macdonell from Glasgow statistics. They are all considerably lower than the correlation between vaccination and severity. The difference would be expected, if the normal distribution held for the total population admitted to hospital. For the tables from which the three former correlations have been calculated are "curtailed" by the omission of the unvaccinated group.

7. The figures from the London epidemic do not enable us to decide whether the correlations of severity with number and foveation of scars are secondary to that between severity and area of scars; though they may be so.

8. The gross correlation between vaccination and severity is lower than the mean net correlation at various ages, owing to the very high correlation between age and vaccination among inmates of smallpox hospitals.

In conclusion, I have to thank Prof. Pearson, of University College, and Dr Ricketts, Superintendent of the Smallpox Hospitals of the Metropolitan Asylums Board, for much help in the preparation of this paper.

* No examples of contingency formulae are given in the paper, as none of the tables were adapted to this method.

MISCELLANEA.

I. On the Distribution of Severity of Attack in Cases of Smallpox.

By KARL PEARSON, F.R.S.

Dr Turner finds that the protection provided by vaccination does not diminish as rapidly with age as he would expect it to do. His expectation arises from the fact that current medical opinion considers that the immunity provided by vaccination diminishes rather rapidly with the increase of the period which has elapsed since vaccination. I do not propose to consider whether protection against death when disease is incurred is really of the same character as immunity against an attack. But I should wish to point out that Dr Brownlee has reached a very similar conclusion to Dr Macdonell on this point by a very different process, and further that it follows, if we simply use the coefficient of association, which makes no appeal to normality of distribution. On the other hand, while I am distinctly interested in Dr Turner's theory of a "curtailed normal curve," I am compelled to say that I consider it an extremely improbable hypothesis in the present application. The group of persons who catch smallpox are a *selected* group of the general population, selected because (i) they have come in contact with the disease, (ii) they were at the time of such contact not sufficiently immune against it. This sufficiency of immunity must depend not only on prior vaccination, but on a host of other causes, the virulence of the poison they encountered, their particular state of health at the time in question, their conduct before and after the risk was run, etc., etc. It is impossible to suppose a rigid line drawn at a certain grade and say all below this grade escape this disease, all above will contract it. I cannot understand how those 'selected' to incur the disease differ from any other naturally selected group with which we are acquainted. Now biometricians are dealing with selected groups every time they measure the variation of a character in a species, but no such truncated normal distributions have yet exhibited themselves*.

Dr Turner says "A curtailed table in measurable characters would result if we collected statistics of height in soldiers, a population from whom all below a certain standard height had already been rejected" (p. 497 above). Now this is a case which can be well tested. For example, height standards exist for both the American and Italian armies. Yet what do we find? That the distribution of the statures of the accepted recruits in both countries, so far from forming a curtailed normal distribution is in some cases as close to a complete Gaussian

* American trotting horses give the nearest approach, there being a time limit to entry in the record of trotters. But this is a perfectly arbitrary line drawn across the trotting population by the hand of man, and not a selection due to a complex of natural causes.

curve as anything with which we are acquainted*. The fact is that there are so many other factors on which selection depends that we get no "curtailing" of the distribution at all. I should imagine that it would be precisely the same with those who are selected to incur small-pox. The selection depends on so very much else than a certain absolutely fixed grade of immunity. This point is of such importance that I think it worth illustrating by an example. It is one which Professor Weldon has most kindly provided me with out of the extensive reductions he has made of Italian conscripts and recruits. In the Province of Verona in the five years between 1875 and 1879 there were 16,203 conscripts† and from these were selected 3810 recruits‡. The following table contains the two distributions:

Stature in Centimetres	*124—125*	*126—127*	*128—129*	*130—131*	*132—133*	*134—135*	*136—137*	*138—139*	*140—141*	*142—143*	*144—145*	*146—147*	*148—149*	*150—151*	*152—153*	*154—155*	*156—157*	*158—159*	*160—161*
Conscripts	4	—	2	1	1	1	—	2	7	7	22	33	55	117	225	626	725	1027	1503
Selected Recruits	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	16	148	239	354

Stature in Centimetres	*162—163*	*164—165*	*166—167*	*168—169*	*170—171*	*172—173*	*174—175*	*176—177*	*178—179*	*180—181*	*182—183*	*184—185*	*186—187*	*188—189*	*190—191*	*192—193*	Totals
Conscripts	1869	2065	2125	1703	1525	1058	699	384	205	131	45	24	7	4	1	—	16203
Selected Recruits	471	542	538	425	414	256	170	133	55	35	9	3	—	2	—	—	3810

In the diagram the two distributions with their corresponding curves are given, and we see that in the selected recruits there is not the slightest approach to a curtailed normal curve. In fact, if we examine the fundamental constants of the distribution§ we find:

For the Conscripts	*For the Recruits*
$\sqrt{\beta_1} = \cdot 1658 \pm \cdot 0130$	$\sqrt{\beta_1} = \cdot 3409 \pm \cdot 0268$
$\beta_2 = 3 \cdot 7516 \pm \cdot 0260$	$\beta_2 = 2 \cdot 7989 \pm \cdot 0536.$

The probable errors are those which would arise if the distribution were truly normal¶. The deviations from symmetry are for the two cases 12·6 and 12·7 times their probable errors respectively. It is impossible therefore to say more than that the two curves are both markedly skew and deviate equally from normality in this respect. In the next place the deviations from meso-kurtosis are in the two cases 28·9 and 3·8 times their probable errors; in other words the curve for the conscripts diverges indefinitely more from the normal curve than that for the selected recruits. This is only one instance out of many which emphasise the same points, i.e. that there is no approach in the selected curve to curtailment and it differs on the whole less from normality than is the case with the unselected material.

* See, for example, Baxter's statistics for U.S. recruits discussed, *Phil. Trans.* Vol. 186, A, p. 385.

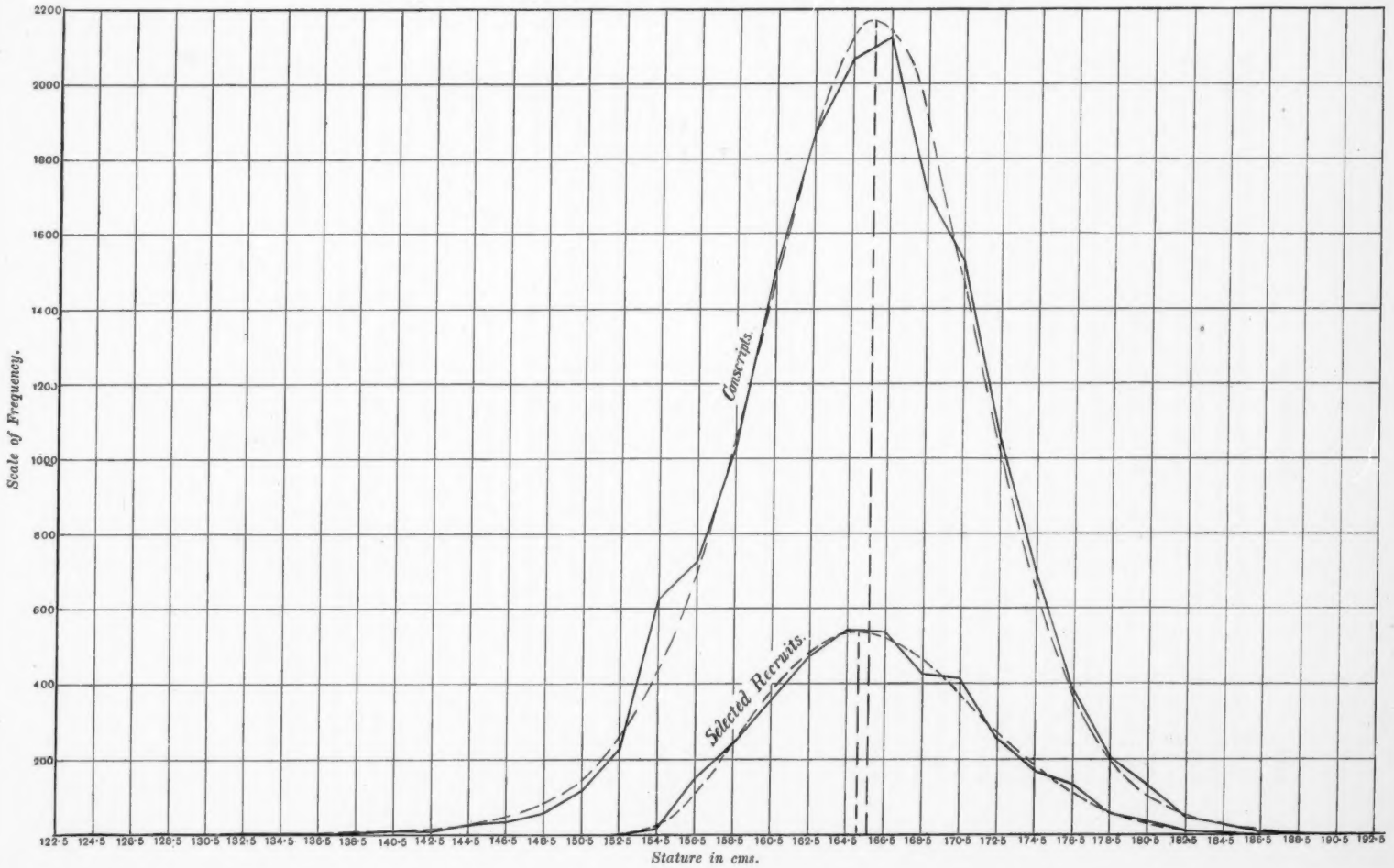
† R. Livi, "Classificazione delle statura.....dei coscritti delle leve di terra negli anni 1875-79." *Annali di Statistica*, Vol. VIII. 1883, pp. 144—9.

‡ R. Livi, *Antropometria Militare*, Parte I., p. 260, Roma, 1898.

§ *Phil. Trans.* Vol. 186, A, p. 368. *Biometrika*, Vol. IV., pp. 174—7.

¶ *Phil. Trans.* Vol. 198, A, p. 278.

Comparison of Frequency Distributions of Verona Conscripts and Selected Recruits.



I have next endeavoured to form an estimate of the distribution of severity in the cases of smallpox, and the only method of doing this seemed to be to form a frequency distribution for the intervals which elapse between either (*a*) onset or (*b*) eruption and the first bath. This may be taken in a rough sort of way as a measure of the severity of the attack. Unluckily any character which depends on time usually gives a skew frequency distribution and the present case is no exception to the rule. But I do not think that the resulting distributions can be in any way described as curtailed normal curves. Dr John Brownlee kindly provided me with the particulars of between 800 and 900 cases. I could distinguish no sensible differences in the male and female distributions. Further, the vaccinated cases formed a very large proportion of the total, for example 779, as against 55 unvaccinated and 21 doubtful. Hence, taking into account that what Dr Turner is dealing with is the distribution of cases in all classes, vaccinated or unvaccinated, I have clubbed all groups together to get my distribution of severity. There were 57 deaths, which must of course be excluded, from a time to bath test of severity, they occurred with rather an erratic distribution at a mean interval of 10·4 days after onset or 7·6 days after beginning of the eruption. Thus the interval between onset and eruption in the fatal cases is about 2·8 days, while in the recoveries it is 3·1 days, so that there may, when more material is forthcoming, be found to be a sensible difference in this interval for the two classes of cases.

The following table gives the distribution of severity of attack as measured by the two tests of days (i) from onset to bath and (ii) from eruption to bath.

Days	*4—5*	*6—7*	*8—9*	*10—11*	*12—13*	*14—15*	*16—17*	*18—19*	*20—21*	*22—23*	*24—25*	*26—27*	*28—29*	*30—31*	*32—33*	*34—35*	*36—37*	*38—39*	*40—41*	*42—43*	Totals
Onset to Bath ...	2	13	40	131	192	152	99	73	40	24	13	17	10	6	4	6	2	1	—	1	826
Eruption to Bath	10	55	164	174	156	96	71	39	26	13	21	6	8	2	7	3	3	1	—	—	855

Neither of these distributions can be considered as a curtailed normal curve. They must, I think, be looked upon as significantly skew distributions of the usual type such as in practice almost invariably arise, when time is the variate to which we plot frequency. They do not appear to me to give any support to Dr Turner's view expressed on p. 496 above as (2), or to the suggestion on p. 497, that the mildest cases are more frequent than those more severe; the different degrees of severity do not diminish without exception as the degree increases, and since the bulk of our cases are all vaccinated, it is not possible to suppose that the actual frequency curve among the vaccinated is of the type suggested by Dr Turner. The modal frequency corresponds to a sensible degree of severity of attack, i.e. to about 13 days from onset to first bath, while the mildest cases correspond to only 4 days. It seems therefore quite impossible to suppose, as Dr Turner does on p. 502, that the severity frequency distribution is half a normal curve of all exposed to risk. The distribution of the attacked very considerably passes the mode, and on the assumption made by Dr Turner, those exposed to risk and escaping must be a very small fraction indeed of the total population exposed to risk. Indeed the above distributions show that the severity of the attack rises from zero to a maximum, and then falls, in the usual skew frequency fashion, at a slower rate to zero. The constants of the above distributions have been worked out; it will suffice here to give, however:

Onset to Bath $\beta_1 = 2{\cdot}3229$, $\beta_2 = 6{\cdot}2466$.

Eruption to Bath $\beta_1 = 2{\cdot}606$, $\beta_2 = 6{\cdot}3664$.

Thus both distributions are sensibly non-Gaussian.

It may be said that this skewness might itself account for the results which Dr Turner considers opposed to current medical opinion and which he would account for by a curtailed normal distribution. Possibly it may, but I hold that this is very unlikely, and for the following reasons:

(i) The severity is here measured by a *time* test, which invariably gives a skew distribution; but if severity could be measured in some other quantitative manner, for example by intensity of eruption, mean rise of temperature or in some such fashion, we should most probably reach a more normal distribution.

(ii) The effect of even the large amount of skewness indicated above on the correlation found by a fourfold division is, I think, not likely to be very large. It is too often assumed that the distribution must be *very closely* normal to give a good result by the fourfold table method. As a matter of fact, if the fourfold division falls on the "long tailed side" of the mode, as it does in these vaccination statistics, we get close values to the actual correlation even by the fourfold table method. To illustrate this I took the only correlation which was available on the data, namely the correlation between onset to eruption and eruption to bath. This correlation was found by the ordinary product moment process, which is independent of any hypothesis of normal distribution and came out $= -\cdot174 \pm \cdot023$. In other words, if severity of the disease be measured by a long period from eruption to bath, then a short period between onset and eruption is associated with severity. This is in accordance with the previous indication that the fatal cases have a short period from onset to eruption.

I now formed a fourfold table from my exceedingly skew distribution, dividing between 14 and 15 days from eruption to bath and between 3 and 4 days from onset to eruption. This gave:

Onset to Eruption.

Eruption to Bath.	*3* days and less	*4* days and more	Totals
14 days and less ...	395	235	630
15 days and more ...	138	51	189
Totals ...	533	286	819

whence the correlation $-\cdot166$ was deduced. It is clear that the skewness has not sensibly influenced the value of the correlation as determined from this fourfold table.

The divergence of the fourfold dividing lines from the median is certainly not quite as great in this case as in the vaccination problem, but it is difficult to go nearer the tails and get anything like a reliable result. Dividing between 4 and 5 days from onset to eruption, and 16 and 17 days from eruption to bath, I get:

Onset to Eruption.

Eruption to Bath.	*4* days and less	*5* days and more	Totals
16 days and less ...	595	104	699
17 days and more ...	112	8	120
Totals ...	707	112	819

In this case the probable error of the frequency 8 is no less than 2, but a change of 2 in the value of this frequency would change the value of r by about 20 p.c. Such a division is

therefore very unreliable*. We find, however, that the correlation $= -\cdot238$ with a probable error of about $\pm\cdot05$. In other words within the limits defined by the probable errors, we get a correlation sensibly equal to the actual value $-\cdot174$. I think accordingly that we may assume that the skewness of our severity test would not materially alter the condition of affairs.

I conclude as follows:

(i) There is no reason *à priori* for supposing that the distribution of severity of attack in smallpox follows a curtailed normal curve. Such curves are contrary to any existing experience of the distribution of frequency in *selected* groups. In particular they do not occur in the case of soldiers suggested by Dr Turner.

(ii) The only tests I have been able to apply to the frequency distribution of severity is that of length of period from onset or eruption to bath. These both give a continuous and not a curtailed curve. The curves are skew, as we are accustomed almost invariably to find them when time is the variate.

(iii) This skewness, due to the time as variate, may not really be characteristic of the distribution of smallpox severity; this could only be determined if we chose a variate other than time to measure severity. But if it were characteristic, it does not appear that a fourfold division taken so far towards the tail as occurs when we divide severity at death would seriously affect the result. It is shewn that within the limits of probable error the product moment method of calculating the coefficient of correlation and the fourfold division method lead to sensibly the same results, even when we have a skewness in the distribution as great as is indicated in the eruption to bath test of severity.

II. On the Mean Duration of Life of Individuals Dying within a Year after Birth†.

By RAYMOND PEARL, Ph.D.

The usual custom in tabulating census returns of mortality at different ages is to use a five year base unit for ages above 5, and below that age a one-year unit. This method of tabulation makes the finding of the moments of the frequency distribution somewhat less simple than would be the case if all the base elements were equal. Furthermore, the age distribution of the heavy mortality of the first year of life is not given at all. It becomes a very important matter to know, at least approximately, this first year distribution when one attempts to find the moments for the whole material, because the frequency in this element must be centred at some point before one can proceed with the calculations. The mean age at death of those dying under one year must be known. In connection with some work on vital statistics which is being carried on in the Zoological Laboratory of the University of Michigan, it became necessary to have as exact a determination of the centering point of this first year mortality as it was possible to get. It is the purpose of this note to present the results obtained.

* Dr Turner's Tables XIV. and XV. contain only 8 and 2 cases in their fourth quadrants and the probable errors of the resulting coefficients are very considerable. I think Dr Brownlee's figures (*Biometrika*, Vol. IV. p. 325; *Journal of Hygiene*, Vol. V. pp. 533—4) may be safely taken as giving a reasonable measure of the reduction with time of vaccination protection.

† Contributions from the Zoological Laboratory, University of Michigan, No. 88.

TABLE I. *Deaths (excluding still-born) at the designated age per* 1000 *legitimate or illegitimate children born.*

Age in days and months	LEGITIMATE					
	1877—1881 average		1882—1890 average		1891—1895 average	
	♂	♀	♂	♀	♂	♀
0 to *1* day	9·32	7·16	10·27	7·89	10·83	8·12
1 „ *2* days	5·41	4·12	5·60	4·30	5·52	4·15
2 „ *3* „	3·52	2·80	3·60	2·77	3·45	2·67
3 „ *4* „	2·33	1·87	2·37	1·85	2·24	1·73
4 „ *5* „	1·93	1·51	1·87	1·60	1·73	1·38
5 „ *6* „	2·32	1·76	2·13	1·64	1·86	1·40
6 „ *7* „	2·87	2·22	2·54	1·89	2·13	1·61
7 „ *8* „	2·71	2·08	2·37	1·88	2·03	1·59
8 „ *9* „	2·31	1·84	2·05	1·62	1·76	1·41
9 „ *10* „	1·86	1·51	1·75	1·38	1·56	1·24
10 „ *11* „	1·90	1·54	1·81	1·46	1·70	1·33
11 „ *12* „	1·89	1·47	1·82	1·43	1·67	1·39
12 „ *13* „	1·96	1·61	1·91	1·54	1·83	1·46
13 „ *14* „	1·93	1·60	1·86	1·54	1·77	1·49
14 „ *1* month*	21·73	18·90	21·97	18·98	22·14	18·80
1 „ *2* months*	22·59	19·76	22·70	19·79	23·20	19·78
2 „ *3* „	18·58	15·86	18·92	16·44	19·72	16·62
3 „ *4* „	15·96	13·59	16·45	14·18	16·90	14·40
4 „ *5* „	13·30	11·26	13·99	11·93	14·15	12·13
5 „ *6* „	11·51	9·85	12·07	10·46	12·24	10·55
6 „ *7* „	10·61	9·05	11·05	9·60	11·01	9·53
7 „ *8* „	9·30	8·33	9·91	8·87	9·72	8·77
8 „ *9* „	8·74	7·96	9·22	8·47	9·18	8·20
9 „ *10* „	8·29	7·83	8·69	8·20	8·30	7·85
10 „ *11* „	7·51	7·22	7·80	7·51	7·44	6·89
11 „ *12* „	6·94	6·92	7·17	6·98	6·69	6·49
	197·32	169·62	201·89	174·20	200·77	170·98
	ILLEGITIMATE					
0 to *1* day	11·48	9·37	12·18	9·62	12·82	10·27
1 „ *2* days	8·10	6·37	8·05	6·23	8·06	6·24
2 „ *3* „	5·28	4·06	5·50	4·32	5·54	4·25
3 „ *4* „	3·54	2·96	3·79	3·14	3·69	2·76
4 „ *5* „	3·32	2·77	3·37	2·85	3·08	2·48
5 „ *6* „	4·09	2·97	3·65	2·85	3·35	2·48
6 „ *7* „	4·54	3·58	4·03	3·29	3·51	2·65
7 „ *8* „	4·06	3·43	3·59	3·17	3·26	2·98
8 „ *9* „	3·62	3·12	3·47	2·64	3·08	2·41
9 „ *10* „	3·28	2·60	3·10	2·54	2·91	2·43
10 „ *11* „	3·53	2·67	3·32	2·67	3·04	2·51
11 „ *12* „	3·30	2·67	3·32	2·94	3·35	2·71
12 „ *13* „	3·43	3·11	3·50	2·85	3·48	2·86
13 „ *14* „	3·33	3·12	3·37	2·84	3·75	2·95
14 „ *1* month*	44·83	39·59	46·19	40·05	48·11	42·00
1 „ *2* months*	53·39	46·48	54·03	47·77	55·03	48·96
2 „ *3* „	40·95	37·06	40·93	37·86	43·36	38·63
3 „ *4* „	31·98	29·74	33·01	30·15	33·84	30·66
4 „ *5* „	24·21	23·04	25·77	23·81	26·49	25·12
5 „ *6* „	19·61	17·97	20·12	19·22	20·49	19·96
6 „ *7* „	16·03	15·38	17·04	16·45	17·45	16·84
7 „ *8* „	12·91	13·00	13·93	13·54	14·04	13·36
8 „ *9* „	11·39	11·11	11·77	12·02	11·51	11·98
9 „ *10* „	9·35	10·00	10·55	10·49	10·33	10·06
10 „ *11* „	8·50	8·63	8·56	8·87	8·27	8·29
11 „ *12* „	7·11	7·37	7·51	8·11	7·39	7·08
	345·16	312·17	353·65	320·29	359·23	322·92

* It is assumed that these are calendar months.

Now it is evident that the larger the number of frequency elements for the mortality of the first year for which we can get data, the more accurate will be the determination of the mean, for the reason that the rate of infantile mortality changes rapidly with lapse of time after birth. It is fairly easy to get data giving the number of deaths occurring in each month of the first year, but a still finer division is desirable. Especially is this true for the first month of life. Roughly from 25 to 45 per cent. of all the deaths falling within the first year occur in the first month. Of these the larger proportion occur early in the month. Fortunately the German statistics, which are in many respects among the best vital statistics available, give a rather detailed age grouping of the deaths under one year. I have used in this work the Prussian statistics* for the years 1877 to 1895, split up into three groups as follows: 1877–81, 1882–90, 1891–95. The data are given in the form of *death-rates per* 1000 *born* of the same class, in the given period. Separate returns are given for male and female, and legitimate and illegitimate mortality. Up to the fourteenth day after birth the mortality for each day is recorded. The mortality of the remainder of the first month is grouped together in a single class, and from the beginning of the second month to the end of the first year after birth the rates are given by months. Still-births are, of course, excluded. A glance at Table I. will make clear the way in which the returns are sub-divided.

It is evident that the grouping here is sufficiently fine to make possible a very accurate determination of the mean age at death. The material was dealt with in the following way: the rates were treated as frequencies, except in the determination of probable errors where, of course, the absolute number of deaths was used. A standard month of 30 days was assumed: then with a unit of 30 days the first and second moment coefficients about an arbitrary axis were determined. From these the position of the mean and the value of the second moment about it were easily found. Only the "rough" second moment was calculated, as it was deemed sufficiently accurate for present purposes, and furthermore it was difficult to determine the proper corrective terms to apply in this case. In the calculations each frequency element was for practical convenience, centred at the mid-point of its range. The error made by so doing is negligible.

The results are shown in Table II.

TABLE II.

Mean Duration of Life in Days of Individuals Dying in the First Year after Birth.

		1877—81		1882—90		1891—95	
		Mean	S. D.	Mean	S. D.	Mean	S. D.
Legitimate	♂	111·44	103·92	113·33	101·13	112·54	100·26
,,	♀	116·75	106·13	118·65	105·97	117·91	104·50
Illegitimate	♂	96·81	91·80	98·64	92·47	98·08	91·41
,,	♀	102·86	94·19	104·94	94·89	103·64	92·92

On account of a lack of necessary data it was not possible to determine the probable errors of all these constants. The general order of magnitude of the probable errors, however, may be seen from a consideration of a single case. For the period 1882–90 data were available from

* *Statistisches Handbuch für den preussischen Staat*, Bde. I. II. and III. Berlin.

which the probable errors could be determined. I find for the probable error of the mean legitimate male mortality 1882–90 a value of ±·0676 day, and for the mean legitimate female mortality ±·0784 day. For the illegitimates the probable errors would be of course somewhat higher, although as nearly as I am able to judge from a rough estimate the probable errors of the means for this group would not exceed ±·15 day. The probable errors of the standard deviations will in all cases be lower. In calculating probable errors the actual number of deaths in the given group and period was taken as n.

A number of points of general interest regarding infantile mortality are brought out by this table. Before considering the main question for which the work was undertaken some of these may be discussed.

(*a*) The mean duration of life is uniformly greater in the case of the legitimates than in the case of the illegitimates. The legitimate males have an average excess of 14·59 days, and the legitimate females an average excess of 13·96 days. The excess is almost exactly the same for each of the three periods, although the mean changes. This uniformity is remarkable, and indicates to what an extent those differences in "nurture" (both pre- and post-natal) to which the difference in mean duration of life of legitimate and illegitimate infants must be attributed are uniform in long periods of time. Associated with the low mean duration of life in the case of the illegitimate infants there is of course a high death-rate as compared with the legitimate group.

(*b*) The illegitimate infants are markedly less variable with respect to duration of life (as indicated by the standard deviation) than the legitimate. The average difference in the case of the males is 9·88 days, and in the case of the females it is 11·53 days. To adopt the illustration which has been used by Pearson, it may be said that the marksman, Death, shoots faster and with deadlier aim at illegitimate than at legitimate infants. This lower variability in the case of the illegitimates may conceivably be the result of a more sharply selective mortality than in the case of the legitimate infants.

(*c*) The mean duration of life of those dying within a year after birth is greater in the two later periods considered than in the first. Apparently, between 1877 and 1896 there has been a gain of about a day. Does this represent a real evolutionary tendency, or is it merely a chance fluctuation? In order to get light on this question I dealt with the mortality of each of the following years separately; 1882, '83, '84, '85, '86, '89, 1890, '91, '94, '95 and '96. These were all the years for which I could get separate data. The groups of male and female, legitimate and illegitimate were treated separately as in the other cases. The results obtained were very interesting in many particulars, but as they fall outside the scope of this note I shall not consider them in detail here. The general tendency from about 1884 on is for the mean duration of life to decrease, with considerable fluctuations from year to year. Thus in 1886 and in 1890 there was an unusually high duration of life. These two years explain why the 1882–1890 group in Table II. shows such high means. In general, a study of these individual year data makes it very clear that there is no steady tendency towards lengthening of the mean duration of life of infants under 1, *within* the period under consideration.

Another interesting point brought out by the single year records is that in the case of both legitimate and illegitimate infants, there is a definitely marked tendency for an increase in the mean duration of life in any year to be associated with an increase in the death-rate for that year. This appears to indicate that in general there is a tendency for any increase in the infantile death-rate to be the result of an increased number of deaths of older rather than younger infants (under 1). It will be noted, however, that this positive relation between death-rate and mean duration of life which appears *within* both legitimate and illegitimate groups is exactly reversed when the two groups are themselves compared. Thus in the legitimate group as a whole we have a condition of low death-rate and high mean duration of life, while in the illegitimate group as a whole the opposite condition obtains.

The single year means are given in Table III.

TABLE III.

Mean Age in Days of Individuals dying under 1 *in the Designated Years. (The death-rate for the same group is given in brackets below each mean age value.)*

Year	♂		♀	
	Legitimate	Illegitimate	Legitimate	Illegitimate
1882	111·71 (202·57)	96·39 (341·90)	117·89 (173·02)	104·17 (314·21)
1883	112·91 (204·94)	100·04 (356·52)	117·72 (175·92)	103·80 (322·44)
1884	114·07 (205·96)	98·86 (360·83)	119·70 (178·13)	103·09 (322·02)
1885	113·91 (197·55)	98·15 (341·24)	119·34 (169·97)	106·26 (311·06)
1886	117·15 (217·29)	101·79 (377·47)	121·46 (188·98)	108·91 (343·06)
1889	112·97 (200·00)	98·00 (363·40)	118·28 (172·90)	104·72 (330·00)
1890	114·62 (204·92)	100·53 (360·65)	122·99 (177·27)	109·05 (324·83)
1891	110·96 (196·55)	96·37 (356·66)	116·26 (166·29)	102·54 (321·75)
1894	113·61 (190·80)	98·15 (339·52)	119·23 (163·65)	104·16 (304·46)
1895	111·89 (207·18)	98·58 (376·59)	116·89 (175·36)	104·07 (337·16)
1896	111·67 (187·62)	95·85 (328·69)	116·83 (159·29)	101·79 (296·99)

(*d*) The mean duration of life is uniformly greater in the females than in the males. The real basis for the difference is not clear. It can hardly be due to differing conditions of "nurture" because there is apparently no reason to suppose that in the long run the environment of a male infant differs in any marked and constant way from that of a female infant during the first year after birth. Along with the lowering of the mean age at death in the male there is a smaller amount of variation in this character.

We may turn now to the practical question as to where the first year mortality shall be centred. The values of the mean for the different groups range between 3 and 4 months and each may be taken as the centering point of its own group. It is desirable, however, to have a single value for the total first year mortality, when males and females, and legitimates and illegitimates are grouped together. In order to obtain such a value resort was had to the plan of a weighted average of all the individual values. Several schemes of weighting were tried, but as all gave very closely accordant results, and the one to be described seemed the most logical, it alone need be detailed here. Examination of all the available data indicated that the absolute number of deaths of males under 1 stood in about the ratio of 1·25 : 1 to the absolute number of

deaths of females under 1. So then the mean age at death of males and females together for each period was determined, the two contributory means being weighted in this proportion 1·25 to 1. Legitimates and illegitimates were dealt with separately. Then the mean age for the whole period 1877–95 for both legitimates and illegitimates was determined. In getting these values the three periods 1877–81, 1882–90, 1891–95 were weighted in the proportions 1 : 2 : 1 respectively. In this way was obtained the mean duration of life of legitimates and illegitimates irrespective of sex. It remained to combine into a single average the legitimates and illegitimates. I found that during the period under discussion there died, on the average, 6·556 legitimate infants under 1, to one illegitimate infant. Averaging, then, the legitimate and illegitimate means previously obtained, with weights of 6·556 and 1, respectively, I found for the mean duration of life of all infants dying under age 1, during the period 1877–95 in Prussia, a value of 113·14 days. With a standard month of 30 days this equals 3·771 months, or ·3142 year. Reckoning 365 days to the year we have from the days directly the mean equal to ·30997 year.

I conclude that the deaths occurring in the first year of life may be centred at ·3 year with sufficient accuracy for ordinary purposes*.

We may turn now to two other practical problems which arise in connection with infantile mortality. These are

1. At what age shall the deaths occurring in the first month of life be centred?

2. At what age shall the deaths recorded in census returns as falling in the age period 0–5 years be centred?

For a solution of the first problem I have taken the Prussian statistics for deaths in the first month of life, during the period 1882–90 (Table I.); I find the following values for the means:

Age at Death of those Dying before completing the First Month of Life.

Legitimate		Illegitimate	
♂	♀	♂	♀
10·82 days	11·25 days	12·41 days	12·75 days

From these results it seems reasonable to conclude that the deaths during the first month of life may be centred, with sufficient accuracy for practical purposes, at ·3 month.

In order to test this result I have calculated the mean age at death of English infants dying under age 1, from the following statistics given in Newsholme's "Vital Statistics†."

Annual Death-rate per 1000 *Living at each Month of Age.*
Healthy Districts.

Age in Months	Death-rate	Age in Months	Death-rate
0	447·51	*6*	70·54
1	145·49	*7*	65·97
2	102·05	*8*	61·85
3	87·16	*9*	58·32
4	81·09	*10*	55·28
5	75·54	*11*	52·86

* The error made by centering each element of the frequency distributions of Table I. at the mid-point would tend to raise slightly the mean age at death for the year. Hence it is probable that ·3 year comes closer to the true value than ·31 year.

† Second Edition, p. 105.

Centering the mortality of the first month at ·3 month and that of each succeeding month at the mid-point I find for the mean 113·09 days, a result in very close agreement with that obtained from the German statistics.

For determining the mean duration of life of those dying within the age class "0–5 years" I have taken as a basis for computation the age distribution of the deaths falling in the first five years of life from the Report of the Twelfth Census of the United States, for the Registration Area. The distribution is as follows:

Age	Under 1	1	2	3	4	Total
Deaths	102220	25986	12020	7825	5520	153571

Now, evidently, the mean age at death for these five groups will fall between

$$\frac{102220\times\cdot3+25986\times1\cdot3+12020\times2\cdot3+7825\times3\cdot3+5520\times4\cdot3}{153571}=\cdot9224,$$

and

$$\frac{102220\times\cdot3+25986\times1\cdot5+12020\times2\cdot5+7825\times3\cdot5+5520\times4\cdot5}{153571}=\cdot9893,$$

if we assume essentially the same distribution of deaths under 1 year in the American population as is shown in the Prussian*. The mean of these two values is ·956 year. The true value probably lies somewhere between ·95 year and ·99 year. With a sufficient degree of accuracy for most cases the deaths recorded in the class "Under 5" may be centred at 1 year.

Briefly stated the results are, then, as follows:

(*a*) The deaths recorded in the age class "Under 1 month" may in practical work be centred at ·3 month.

(*b*) Those recorded in the age class "Under 1 year" may be centred at ·3 year.

(*c*) Those recorded as "Under 5 years" may be centred at 1 year.

* Since the deaths in the second, third, fourth and fifth years of life clearly will not centre lower than those of the first year, nor higher than the mid-point of each year.

CAMBRIDGE: PRINTED BY JOHN CLAY, M.A. AT THE UNIVERSITY PRESS.